DOCUMENT RESUME

ED 203 199 CE 029 396

AUTHOR	Pritchard, Robert D.: And Others
TITLE	Enhancing Productivity Through Feedback and Goal Setting. Final Report.
INSTITUTION	Houston Univ., Tex. Dept. of Psychology.
SPONS AGENCY	Air Force Human Resources Lab., Brooks AFB, Tex. Occupational and Manpower Research Div.
REPORT NO	AFHRL-TR-81-7
PUB DATE	Jul 81
CONTRACT	F33615-78-C-0025
NOTE	61p.: Tables may not reproduce well due to small print.
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Adults; *Employee Attitudes; *Feedback; Field Tests; *Goal Orientation; Influences; *Job Performance; Motivation; Motivation Techniques; *Productivity; *Work Attitudes
IDENTIFIERS	*Goal Setting

ABSTRACT

A field test was conducted to research the effects of feedback and goal-setting techniques on increasing productivity. Subjects were regular employees of two autonomous clerical-type units in a credit card and payment processing center of a Southwestern oil company. The study design had three phases--baseline period and two experimental conditions in which different combinations of feedback and goal-setting procedures were instituted. Attitude and performance data were collected in each of the three phases. The results indicated that both feedback and goal setting produced meaningful increase in productivity. Other major findings were (1) personal feedback was as effective as impersonal feedback, (2) absolute feedback was as effective as comparative feedback, (3) goal setting with feedback was more effective than feedback alone, (4) poorer performers were positively influenced by the treatments while better performers were largely unaffected, and (5) attitudes under feedback and goal setting were as good or better than before the treatments. (Author/YLB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

AFHRL-TR-81-7

LE

AIR FORCE HUMAN RESOURCES LABORATORY

ED203199

ENHANCING PRODUCTIVITY THROUGH FEEDBACK AND GOAL SETTING

By

Robert D. Pritchard
David G. Bigby
Mark Beiting
Susan Coverdale
Carol Morgan
University of Houston
Psychology Department
4800 Calhoun
Houston, Texas 77004

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235

July 1981

Final Report

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Approved for public release; distribution unlimited.

CE029396

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by the University of Houston, Psychology Department, 4800 Calhoun, Houston, Texas 77004, under Contract F33615-78-C-0025, Projects 7734 and 2313, with the Manpower and Personnel Division, Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235. Clessen J. Martin was the Contract Monitor for the Laboratory.

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

NANCY GUINN, Technical Director
Manpower and Personnel Division

RONALD W. TERRY, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFHRL-TR-81-7	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ENHANCING PRODUCTIVITY THROUGH FEEDBACK AND GOAL SETTING		5. TYPE OF REPORT & PERIOD COVERED Final 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Robert D. Pritchard, Susan Coverdale, David G. Bigby, Carol Morgan, Mark Beiting		8. CONTRACT OR GRANT NUMBER(s) F33615-78-C-0025
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Houston, Psychology Department 4800 Calhoun Houston, Texas 77004		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62703F 61102F 77340809 2313T108
11. CONTROLLING OFFICE NAME AND ADDRESS HQ Air Force Human Resources Laboratory (AFSC) Brooks Air Force Base, Texas 78235		12. REPORT DATE July 1981 13. NUMBER OF PAGES 62
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Manpower and Personnel Division Air Force Human Resources Laboratory Brooks Air Force Base, Texas 78235		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) feedback goal setting performance feedback		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The present research represents a field test of the effects of feedback and goal setting techniques on increasing productivity. Absolute and comparative, as well as personal and impersonal, feedback were manipulated. Goal setting was added to several of the feedback groups. Subjects were regular employees working on two clerical type jobs. The results indicated that both feedback and goal setting produced meaningful increases in productivity. Other major findings were (a) personal feedback was as effective as impersonal feedback, (b) absolute feedback was as effective as comparative feedback, (c) goal setting with feedback was more effective than feedback alone, (d) poorer performers were positively influenced by the treatments while better performers were largely unaffected, (e) attitudes under feedback and goal setting were as good or better than before the treatments. The findings are discussed in terms of potential application to ongoing Air Force settings.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SUMMARY

Objective

The objective of this research was to field-test the productivity-enhancing effects of different types of performance feedback techniques developed from previous research. The effects of goal setting as a productivity-enhancing procedure were also investigated.

Background

The present research is the final phase of a four-stage program of productivity research. In the first three phases, the literature was examined, and specific intrinsic motivational variables were tested in a controlled setting. This final phase is concerned with the effects upon productivity of manipulating goal setting and feedback techniques. These two variables were selected on the basis of showing promise from earlier projects and because they could be implemented in an operational Air Force environment. Results from this investigation may be used by Air Force managers to design and implement feedback and goal setting systems for increasing productivity.

Approach

This study was a field experiment conducted in the credit card and payment processing center of a large company headquartered in the Southwest. The design of the study involved the recording of rate and accuracy measures during a baseline period and then instituting feedback techniques during the first treatment phase and feedback plus goal setting techniques during the second treatment phase. Attitude and performance data were collected in each of the three phases.

Specifics

The basic feedback manipulation for the first treatment phase consisted of providing the employees with a computer feedback report on their individual performances. A comparative feedback manipulation involved the reporting of individuals' ranking within their work group on each of the major performance indices. An impersonal feedback manipulation was achieved by having a clerk distribute the reports instead of the supervisor. The personal feedback manipulation involved the reports being reviewed by the supervisor, who distributed them personally to the employee. Goal setting conditions were applied to these feedback conditions in the second treatment phase.

Results showed an overall positive effect on performance. Increases in quantity of output typically ranged from 5% to 10% with a mean increase of 6.4%. Error rates decreased by an average of 11%, with over half of the decreases in the 15% to 28% range. Other major findings were (a) impersonal feedback was as effective as personal feedback, (b) absolute feedback was as effective as comparative feedback, (c) goal setting with feedback was more effective than feedback alone, (d) treatment effects were strongest among employees who were initially low performers and were relatively slight among those who were initially high performers, and (e) employee attitudes under feedback and goal setting were as good or better than before the treatments.

Conclusions

One of the most important conclusions from this study is that feedback systems should be designed so that individuals can observe that increases in effort produce changes in level of performance. Goal setting and feedback should be frequent and objective, and should take individual differences into consideration. Also, separate feedback for each distinct type of job activity should be provided. Finally, this study has shown that goal setting and feedback techniques are relatively simple to apply and can be implemented by local managers.

PREFACE

The work reported in this study was initially funded by the Air Force Office of Scientific Research (AFOSR) under Work Unit 2313T108 and was completed under Work Unit 77340809, "Improved Productivity Through Use of Intrinsic Rewards and Feedback Techniques."

The concern for increased military effectiveness has created a demand for practical productivity programs which can be implemented by local managers. The present project represents a field test of the effects of feedback and goal setting techniques.

Special appreciation is expressed to Dr. William Alley, Dr. Joe T. Hazel, and Dr. Raymond E. Christal who provided sustained interest and support for productivity research. This project was initially monitored by Maj John O. Edwards, Jr., and later by Dr. Clessen J. Martin.

Table of Contents

Introduction 1

Method 10

Results 15

Discussion and Conclusions 45

References 47

Appendix A 51

Appendix B 53

I. Introduction

The enhancement of productivity is an issue of central concern in today's Air Force. The Air Force Human Resources Laboratory has dealt with this issue in an extensive program of research (Christal, 1973; Gould, 1976; Tuttle, Gould, & Hazel, 1975) which has investigated job satisfaction, retention, motivation, rewards, and productivity. One facet of this research program has been to investigate motivational techniques with potential to enhance productivity, with special emphasis on intrinsic motivation techniques.

The logic of this program is that approaches to increasing motivation have traditionally been divided into two major types. The first is extrinsic motivation techniques which make valued, organizationally controlled rewards contingent on effective performance. Examples of such techniques would be piece rate pay systems, bonuses for high performance, and special forms of formal recognition for high performance. While all would agree that extrinsic rewards are necessary for any job, trying to increase productivity through increasing performance-contingent extrinsic rewards has its problems. Such programs are difficult to administer, and to be really effective, the rewards must be fairly large, thus making such programs expensive (Pritchard, DeLeo, & VonBergen, 1974).

The second major approach to increasing motivation is through intrinsic techniques. The goal here is to somehow structure the job so that the person experiences some sort of positive affect when he or she performs well, and negative affect when he or she performs poorly. One example of this positive affect would be a feeling of personal accomplishment for a job well done. If jobs can be structured so that people do get feelings of personal accomplishment from doing a good job, and feelings of personal dissatisfaction when they do a poor job, several benefits accrue. First, people reward themselves. That is, the basic source of motivation is internal rather than external. Second, the motivation so produced is more permanent and does not require an elaborate externally administered program.

Unfortunately, however, while intrinsic motivation techniques have great promise for enhancing productivity, relatively little is known about them. It is only in the last 10 years or so that behavioral scientists have begun a systematic study of this area. This real potential of intrinsic motivation techniques for impacting productivity coupled with a lack of knowledge about them led to the current research program.

Plan of the Research

The present research is the final phase of a four-stage program of research. In the first phase, the existing literature was examined to isolate those variables that had promise for affecting intrinsic motivation. This search (Pritchard & Montagno, 1978) identified fourteen variables:

1. Feelings of personal control over the task.
2. Feelings of competence at doing the task.
3. Contingent extrinsic rewards. (Negatively related)
4. Degree of variety in the skills required to do the task.
5. Degree to which the task required the use of valued abilities.
6. Degree to which the person identifies with the task.
7. Degree to which the person does a complete unit of the task.
8. Perceived significance of the task.
9. Degree of autonomy on the task (related to #1).
10. Adequacy of performance feedback.
11. Higher order need strength.
12. Work values.
13. Cultural influences.
14. Optimal arousal level.

In the second phase, some of these variables were explored in a controlled setting to begin to assess their suitability for eventual field application. Feelings of personal control and competence, as well as contingent extrinsic rewards, were examined by Fisher and Pritchard (1978). Performance feedback was addressed by Pritchard and Montagno (1978).

The third phase attempted to isolate variables which could be implemented in an operational Air Force environment and to test a fairly

large number of different possible applications in a controlled, yet realistic setting. In this stage, it was necessary to narrow the list of potential determinants of intrinsic motivation to a smaller subset for more careful study. After evaluating them in terms of (a) their potential use in a field setting, (b) the feasibility of testing them in the work simulation setting to be used, and (c) the "quality and quantity" of previous literature available, major emphasis was placed on the performance feedback variable. Six dimensions of feedback and a job design variable, completeness of the task unit, were evaluated in the controlled setting. The major conclusions of this study were that feedback had meaningful potential for increasing productivity.

In the fourth and final phase, described in this report, several specific types of performance feedback, singly and in conjunction with goal setting, were selected to be tested in an operational work environment similar to those found in some Air Force settings.

Review of the Relevant Literature

Feedback

Feedback has long been recognized as a key element in learning, based on a number of assumptions about the motivational, learning, and reward properties of feedback (Ammons, 1956; Annett, 1969; Thorndike, 1927). As a result, a considerable body of research on feedback has accumulated over the years. (See Ammons, 1956; and Ilgen, Fisher, and Taylor, 1977, for reviews of this literature.) Unfortunately, as Ilgen and his colleagues point out, despite this plethora of research, not much is known about feedback as a psychological process. This problem arises because of the rather simplistic approach used in applying the feedback variable in many laboratory experiments. The result is that the literature is filled with many studies dealing with only a single dimension of feedback. Some examples of the more popular dimensions looked at are immediate versus delayed feedback (Beeson, 1973; Christian, 1972; Robinson & Kulp, 1970; Sassenwrath & Younge, 1969), knowledge of right versus wrong responses (Longstreth, 1970; Merrill, 1970; Wike, 1970), verbal versus nonverbal (Lair & Smith, 1970; Simpson, 1972), personal versus impersonal (Weidner, 1968), accurate versus inaccurate (Griswold, 1970), information content (Berman, Fraser, & Theious, 1970), and frequency (Ivancevich, Donnelly, & Lyon, 1970).

Feedback has recently come to the attention of industrial/organizational psychologists from two directions. The first of these is the perennial problem faced by organizations in the area of performance appraisal. McCall and DeVries (1976) point out that managers are aware that feedback is necessary for improving performance of employees. Managers are often reluctant, however, to give performance appraisal feedback because of a number of contextual characteristics that can be

a source of conflict between supervisors and subordinates. The authors cite as an example the fact that performance appraisal is intended to be on the individual level, yet the individual's performance is often hopelessly intertwined in a group performance. McCall and DeVries list a set of characteristics that describe what some would consider to be the contemporary "ideal" performance appraisal system. Among these characteristics are objectivity, participation, and frequency. Leskovec (1967) would add to this list a specificity dimension. Porter and Lawler (1968), working from the expectancy-valence model, attempt to explain why feedback should have an effect in organizational settings. They argue that feedback serves the role of establishing a link between effort and performance, and between performance and rewards. To the extent that feedback performs this function, it will be effective.

The second route by which feedback has gained attention is the work done on the motivational aspects of job design. Basing their work on Turner and Lawrence (1965), Hackman and Oldham (1975) developed a list of core job dimensions which should influence the motivating potential of a job. One of these core dimensions was feedback. Hackman and Oldham (1975) include feedback as one of the key elements in their instrument, the Job Diagnostic Survey (JDS). For the purposes of the JDS, feedback is divided into two categories: feedback from the job itself and feedback from agents (i.e., people). Feedback from the job is defined as the degree to which performing a particular job results in the job incumbent's obtaining direct and clear information about the effectiveness of his or her own performance. Feedback from agents is information received from supervisors or coworkers (Hackman & Oldham, 1975). Hackman and Oldham further theorize that the presence of feedback in a job will lead to what they call a critical psychological state, namely, knowledge of results (KR). The distinguishing characteristic of the KR state is that the worker has continual knowledge of his or her performance.

Hackman and Lawler (1971) propose that regardless of what other characteristics a job might have, if there is no way for the individual to get feedback, he or she will be unable to experience higher order satisfaction. Hackman and Lawler (1971) in an early evaluation of this model had some difficulty identifying the feedback content of jobs. They discovered in their analysis that there was considerable lack of inter-rater agreement as to the amount of feedback various jobs actually contained. Four rating procedures were used (employees, supervisors, researchers, and the Turner and Lawrence classification) and no significant correlations were found among them. The results of their study did show, however, that feedback was related to employee reactions to his or her job.

Payne and Hauty (1955) identified what they referred to as the motivational and incentive functions of feedback. In most organizational settings, the major concern is with the motivational facet. In other

words, the concern is for altering the effort level in response to feedback information (Porter & Lawler, 1968); this assumes that no more learning is required to adequately perform the task. Locke, Cartledge, and Koeppel (1968) make a similar distinction. They propose that when no more learning occurs or when the feedback contains summary information only, the motivational process of feedback is in operation. They further contend that knowledge of results is motivational only to the extent that it facilitates the setting of goals. This implies that the feedback must contain information about standards or past performance.

Annett (1969) has proposed that feedback is a multidimensional concept and that simple interpretation of feedback as a facilitator of learning or an enhancer of motivation may be unjustified. He proposes that the effects of feedback depend on two factors. First, the nature of the information must be considered. Issues relevant here would be the various sources of feedback and methods of presentation. The second factor Annett considers important is the use made of the feedback information. The implication here is that any given episode of feedback may have a number of purposes, only one of which may be motivation.

A second problem in trying to isolate the issue of motivation is that of the narrow interpretation of learning. It would appear that learning in organizational settings goes beyond merely mastering a set of tasks. It is conceivable that at any given point in time, a particular employee may be undecided about the particular task in which to be engaged. This implies that the employee's role is dynamic and the selection of tasks is a constant learning process significantly influenced by feedback.

There have been a number of attempts, recently, to arrive at some convergence on the meaning of feedback. Greller and Herold (1975), for example, have studied the "source" dimension of feedback. Using a diverse sample of workers, they established five levels of the source dimension: the company, the supervisor, the coworkers, the task, and the worker's own feelings and ideas. Their results showed that people most often used sources close to themselves for feedback. That is, one's own feelings and the task were the most common sources of feedback. It should be noted that the categories proposed by Hackman and Oldham (1975), namely the task itself and agents, could be looked at as summary headings for the Greller and Herold categories.

Ilgen, Fisher, and Taylor (1977) have also selected "source" of feedback as an important dimension. In addition, they propose that every external source varies in two important characteristics: credibility and power. They argue that the higher a source rates on these two factors, the more positive will be the effect of feedback. Ilgen and his associates also propose that the notion of separate feedback from the task and from the self, while theoretically possible, is probably unlikely in organizational settings.

Ilgen, Fisher, and Taylor (1977), in addition, have derived two other important dimensions of feedback. One of these is a qualitative dimension. This dimension reflects feedback characteristics such as immediacy of feedback, time delay, individual or group, and so on. The final dimension proposed by Ilgen et al. is purposefulness. This refers to the fact that an agent usually has some motive for presenting feedback to an employee.

Greller (1975) noted some of the problems with using simple dichotomies on single dimensions of feedback. He then developed a taxonomy of feedback containing four dimensions. These include consequences from agents, expressions from agents, task feedback, and internal comparison. The internal comparisons dimension reflects the idea that individuals observe others on the same job as themselves and then make internal comparisons. In addition to these four categories, Greller argues that the evaluative nature of the information presented is important, that is, each of these dimensions can vary along a positive-negative continuum.

In a study designed to test this model, Greller (1975) found partial support for his hypothesis. One discrepant finding was that positive and negative feedback were not simple opposites. He concluded that individuals will select the most favorable information from feedback in order to develop the most positive image of themselves under the given condition. One extension of Greller's conclusion is that, to the extent self-enhancement leads to feelings of competence, there could be increases in intrinsic motivation (Deci, 1971).

In another attempt at developing a meaningful taxonomy of feedback dimensions, Greller and Herold (1977) performed a factor analysis on a questionnaire of 50 feedback items. The result was a five-factor solution: Factor I - negative feedback, regardless of source; Factor II - positive, from above; Factor III - positive from nonhierarchical others: Factor IV - internal criteria; and Factor V - work flow feedback. Two overall dimensions that clearly stand out from these data are source and positive-negative.

After reviewing the feedback literature, Pritchard and Montagno (1978) argued that a more extensive taxonomy of feedback dimensions was needed. Based on this review of the literature and an intuitive analysis, they developed a list of 14 comprehensive dimensions. It was acknowledged that in some instances the dimensions overlap to some degree, but it was felt that each had its own key elements and should be separately defined.

Feedback Dimensions

1. Positive vs. Negative. This dimension has three permutations: positive alone, negative alone, or positive and negative presented together. This dimension deals directly with the correctness of the

behavior of interest. If the behavior is correct, positive feedback is given. If it is incorrect, negative feedback is given. If we are dealing with positive alone or negative alone, then feedback speaks only to correct or incorrect behavior, not both.

2. Timing of Feedback. This dimension refers to the time that elapses between the performance of a task and the presentation of feedback. This elapsed time might vary from a long span of months or years, to a situation where feedback is available during, and immediately after, performance.

3. Specificity. Variation along this dimension concerns the molecularity of the behaviors on which the feedback is given. The extremes of specificity would range from a single evaluation of a person's total role, to feedback on the smallest task-relevant act in which the person engages. Other meaningful points along this continuum might be various sub-products or major products the individual produces.

4. Evaluative - Non-Evaluative. Evaluative feedback is feedback given by some powerful other in the organization, and clearly implies that the powerful other has evaluated the performance of the person. Non-evaluative feedback does not include this formalized evaluation by another person. Non-evaluative feedback typically would come from mechanical sources that do not involve another person.

5. Absolute - Comparative. Absolute feedback is information only about a person's own performance. Under comparative conditions, individuals would know their own performance level, as well as how this level of performance compares with reference to some other group, such as their own work group.

6. External - Internal. External feedback is information which comes from a source external to the performer. This source could be another person or some mechanical device (e.g., a counter). Internal feedback refers to information which is based on the person's own experience with the task. Proprioceptive or kinesthetic feedback would be internal types of feedback.

7. Personal - Impersonal. This dimension is concerned with the level of personal contact between the performer and the source of feedback. Face-to-face oral feedback from the supervisor would be highly personal, while a self-obtained computer printout outlining performance would be highly impersonal.

8. Power of Source. Power here is defined in terms of the ability of the source to control the individual's rewards. A high power source would control pay raises, promotion, or social rewards. A low-power source, conversely, would control no rewards.

9. Schedule of Feedback. This dimension basically reflects the reinforcement schedule of the feedback. Examples of such schedules would include continuous (after every response), fixed interval (weekly, yearly), and variable interval (at different points around some average length of time).

10. Group vs. Individual. This dimension concerns whether the feedback presented deals with individuals alone or with the entire work group. For example, information about the progress of a given group project may tell the individuals very little about their own behavior.

11. Comprehensiveness. This dimension is defined as the percentage of the role covered by the feedback. If the feedback dealt with only one aspect of a complex job, it would be low in comprehensiveness.

12. Formal - Informal. Feedback along this dimension concerns whether or not the individual has an expectation of receiving feedback prior to the feedback encounter. An annual performance appraisal interview would be an example of formal feedback. Informal feedback is more random in nature and would not be expected prior to the encounter.

13. Public - Private. This dimension refers to whether feedback is given to the individual alone or in the presence of others. These others would most generally be members of the individual's work group.

14. Accuracy. Accuracy refers to the validity of the information; that is, the extent to which the information given to the person validly reflects the true state or nature of his or her performance.

Two studies have directly tested parts of this feedback taxonomy. Pritchard and Montagno (1978) experimentally tested two dimensions of the feedback taxonomy, specificity and absolute-comparative, in a job simulation experiment. People were hired to work for 6 days at a clerical inventory control job, and the effects of feedback on performance and job satisfaction were assessed. Non-specific feedback had strong positive effects on productivity while specific feedback had no effect. In non-specific conditions, comparative feedback was superior to absolute; in specific conditions, no effects emerged.

Pritchard, Montagno, and Moore (1978) manipulated six of the feedback dimensions during a 3-week job simulation experiment using the same inventory control task. In this study, Pritchard et al. found both personal-evaluative and impersonal-nonevaluative feedback to positively affect performance, with impersonal feedback being superior. Also, both high and low specificity feedback were effective but high specificity feedback was clearly more effective. Group and individual feedback procedures were equally effective. Finally, making feedback public or keeping it private had no differential effect on performance, either positive or negative. There were no appreciable differences in employee

satisfaction on any of the dimensions manipulated.

The best combination of feedback procedures in this study was impersonal, high specificity, individual feedback in either a public or private format. This combination produced a 26% increase in quantity of output and a 27% decrease in errors made.

Based on the results of the Pritchard and Montagno (1978) and the Pritchard, Montagno, and Moore (1978) studies and the operational limitations of the field organization, the personal-impersonal and comparative-absolute dimensions were selected for the primary feedback manipulations. In addition to the manipulated dimensions, the treatments were fixed at constant positions on a number of other feedback dimensions. All conditions were designed to be specific enough for the worker to get information on what behavior to change, but not so specific as to cause information overload. Feedback was individualized and private. Feedback was comprehensive in that it covered as much of the total job duties as possible, and it was formal in that it was expected at regular daily intervals. The source of the feedback was external to the individual; it was both positive and negative in that information would be given on good and poor performance; and it was as accurate as possible.

Goal Setting

The second major set of experimental manipulations dealt with goal setting. Goal setting was selected as a manipulation because it has been shown to be effective in increasing performance in a wide variety of situations (Locke, 1968; Steers & Porter, 1974; Latham & Yukl, 1975) and because of the complementary relationship between feedback and goal setting in many situations.

Locke's (1968) theory of goal setting is concerned with the relationship between conscious goals or intentions and task performance. According to this theory, the conscious intentions of individuals regulate their actions. A goal is defined simply as what the individual consciously intends to do. According to Locke, hard goals produce higher performance than easy goals. Hard goals that are specific are more effective than no goals or than a generalized goal of "do your best." Goals mediate the relationships between performance and the effects of monetary incentives, time limits, knowledge of results, participation in decision making and competition. Further, assigned goals are effective only to the degree to which they are accepted.

Latham and Yukl (1975) in their extensive review of the application of goal setting in organizations concluded that goal-setting programs tend to be "effective over an extended time period in a variety of organizations at both the managerial and non-managerial levels." (p.385) They found strong support for the proposition that specific goals

increase performance and that difficult goals, if accepted, result in better performance than do easy goals. Latham and Yukl found no field study support for the mediation hypotheses. Finally, Latham and Yukl found most studies of participative versus assigned goals to have some evidence of the superiority of participative goal setting, but significant differences were found only under limited circumstances.

Based on the goal setting literature and especially the Latham and Yukl review, the goal setting manipulation was defined as follows. Specific goals were set by an interaction between supervisor and worker with the full participation of the worker. The worker was encouraged to set moderately difficult to difficult goals and to commit himself or herself to achieving those goals.

Thus, the present study attempts to build on the earlier work by taking those feedback procedures that were effective in the laboratory settings and test them in a field setting. Specifically, the research compares personal with impersonal feedback and compares absolute with comparative feedback. In addition, this research examines the effects of goal setting over and above feedback alone.

Method

Overview

This study was a field experiment conducted in the credit card and payment processing center of a large oil company headquartered in the Southwest. The subjects were the regular employees of two autonomous clerical type units: the Remittance Control Section and the Data Input Section. Each of these sections operated two separate shifts.

The design of the study had three phases: first, a baseline period, then two experimental conditions in which different combinations of feedback and goal setting procedures were instituted. Attitude and performance data were collected in each of the three phases.

Description of Jobs

In the Remittance Control section, customers' credit card payments were processed for input into a computerized accounting system. The two major work activities involved (a) opening envelopes, visually scanning the payment check and the payment stub, and sorting the check/stub pairs for later processing, and (b) encoding the bottom of the payment stub and check with the amount paid in magnetic ink. Sorting was done by groups of four employees around a constantly paced opening machine. Encoding occurred on individualized, self-paced machines. Individual accountability was maintained for both activities. Under normal circumstances, workers rotated between the sorting and encoding tasks at approximate 90-minute intervals.

During the course of the study, the number of employees in a treatment varied with the normal, ongoing hiring and termination processes of the section. In total, 96 employees in the day shift and 37 employees in the evening shift were affected at one time or another. In both shifts, employees were predominantly females, with 91 females and 5 males on the day shift and 36 females and 1 male on the evening shift. The average age of all employees was 24.3 with a range of 17 to 63.

In the Data Input section, various items of customer account information were keyed on visual display computer terminals by employees in a key-to-disk operation. The type of information keyed varied considerably but almost always was encompassed by one of approximately 90 standard keying formats. Individual employees typically keyed in only a small subset of these formats on a demand basis. Individual keying rate and keying time information was maintained by format for each employee. In addition, employees were held accountable for the proportion of total available work time spent on keying and other work related activities.

As in Remittance Control, Data Input had a day and an evening shift. A total of 32 employees on the day shift and 24 employees on the evening shift were affected by the treatments at some time during the study. Again, they were predominantly female. On the day shift, there were 32 females and no males, on the evening shift there were 22 females and 2 males. Average age for the entire section was 30.8 with a range of 18 to 56.

Experimental Conditions

As it has been previously stated, there were two experimental phases. In the first, all conditions involved the institution of feedback. In the second, goal setting was added to the existing feedback in all but one condition. In this condition, the type of feedback was changed.

The basic feedback manipulation in all conditions consisted of providing the employee with a computer-generated feedback report on his or her performance. These reports were given to employees daily. In the Data Input section, the report was for the previous day; in the Remittance Control section, there was a 4- to 5-day lag. In general, the feedback report contained the employees' performance on their major performance indices for the report date and their averages for the previous week.

In Remittance Control, employees were given their sorting error rate, their encoding rate on three different types of encoding activities and an overall encoding error rate. In addition, the feedback report included a summary encoding effectiveness index that integrated the output rates for the three different types of encoding and the encoding error rate. In a crude cost analysis, it was determined that the

time cost of correcting one encoding error was five times the cost of encoding one unit correctly. The summary index, therefore, charged the employee five times his or her error rate. The result was an index that was 100% when the employee encoded at the standard rates with no errors and varied higher as encoding rates increased and lower as the encoding error rate increased. (An example of a Remittance Control feedback sheet is provided in Appendix A.)

In Data Input, performance in each format was compared to the standard for that format, and a composite index was formed. The percentage of standard performance in each format was weighted by the time spent in that format to form an overall keying effectiveness index. In addition to their keying effectiveness index, employees were given an index of the percentage of time they spent in keying and other work activities, relative to the total amount of time available for work. Finally, their reports included an overall work effectiveness index which was the product of the keying index and the time index. (An example of a Data Input feedback report is provided in Appendix B.)

For both sections, the basic feedback report was modified for the various experimental conditions. In the comparative conditions, the reports also contained the individual's ranking within the work group on each of the major indices. In the impersonal condition, the relationship between supervision and the feedback reports was minimized. Reports were distributed by a clerk, supervisors did not review the reports, and supervisors did not comment upon their contents unless requested to do so by the employee. In the personal condition, supervisors reviewed the reports, circled very high or very low indices, wrote their initials on them, and passed them out to employees personally. In addition, once a week they wrote short evaluative comments on each employee's report.

In all of the goal setting conditions, each employee met with his or her supervisor four times to establish goals. For the first session, the supervisor was given a set of suggested low, moderate, and difficult goals based on the individual's past performance and a listing of the individual's average weekly performance for the previous 10-12 weeks. The supervisor and subordinate discussed the suggested goals and jointly established a goal for each of the major performance indices (sorting error rate and encoding effectiveness in Remittance Control and keying effectiveness in Data Input). After 2 weeks, they met again, evaluated goal attainment, and either retained the same goals for the next 2 weeks or changed them up or down. This meeting was repeated after 4 weeks and 6 weeks.

Since new employees were regularly being hired and trained in both sections, learning could not be discounted as a factor affecting individual performance. Therefore, suggested goals were established by predicting the level of employees' performance given their level of tenure

at the start of the goal setting treatment. Using baseline data for all individuals within a section, learning curves were established for each of the major indices. Using these curves and the employee's average relative position above or below the curves, predictions of performance were made. To this predicted performance, small, moderate, and large increments were added to produce low, moderate, and difficult goals. Given that it was probable that individuals varied considerably in their capacity for improving the level of their output, the size of the increments used for a particular employee were based on the historical variability of that individual on the given index. The results of this process were suggested low, moderate, and difficult goals that had been individually tailored to the employees based on their position on the learning curve and on their potential for improvement.

Design

The design for the study is shown in Figure 1. As can be seen, each shift was treated as a separate experimental group. Each group received a combination of the two dimensions of feedback during the first treatment. In three of these groups, the day and evening shifts in Remittance Control and the evening shift in Data Input, goal setting was added during the second treatment. In these conditions, the type of feedback remained constant. In the fourth condition, the day shift in Data Input, the type of feedback given was changed, and goal setting was not added.

Operational considerations within the two sections dictated the beginning dates and duration of the treatments. There was considerable variability across sections and some variability within the Data Input section. As can be seen from Figure 1, baseline data in the Remittance Control section were collected for 2 months; the first treatments lasted for approximately 4 months, and the second treatments lasted for 2 months. In Data Input, baseline data were collected for 5 months; the first treatments lasted approximately 6 months, and the second treatments lasted approximately 3 months.

Performance data were collected in both sections from the beginning of the baseline period through the end of the second treatments. For the purposes of the data analysis, three performance variables were tracked in the Remittance Control section: the sorting error rate, a combined encoding rate (in which each of the three types of encoding were weighted by difficulty) and the encoding error rate. (The encoding effectiveness index was ignored because it was redundant with the other two encoding variables.) Only the keying effectiveness variable was considered for the major analyses in Data Input.

In addition to the performance variables, attitude data were collected by the administration of the Minnesota Satisfaction Question-

Section	Shift	Baseline	First Treatment	Second Treatment
Remittance Control	Day Shift	Baseline (2 months)	Personal, Comparative Feedback (4 months)	Personal, Comparative Feedback + Goal Setting (2 months)
	Evening Shift	Baseline (2 months)	Personal, Absolute Feedback (4 months)	Personal, Absolute Feedback + Goal Setting (2 months)
Data Input	Day Shift	Baseline (5 months)	Impersonal, Comparative Feedback (6 months)	Personal, Comparative Feedback (3 months)
	Evening Shift	Baseline (5 months)	Personal, Comparative Feedback (6 months)	Personal, Comparative Feedback + Goal Setting (3 months)

Figure 1 The Experimental Design.

naire (Weiss, Dawis, England, & Lofquist, 1967) near the end of the baseline periods, the end of the first treatments, and again near the end of the second treatments.

Procedures

At the beginning of the feedback treatments, the feedback program and the computer generated feedback reports were carefully explained to the employees and their supervisors in small group sessions. These orientation sessions included an explanation of each item on their feedback sheets and a general discussion of example reports.

For the goal setting treatments, supervisors were carefully trained. Emphasis was placed on employee participation in the process, the setting of specific, moderate to difficult goals, and employee acceptance of the goals. Goal setting sessions were modeled for the supervisors, then they role-played the process, alternatively acting the role of the supervisor and role of the employee, until each supervisor understood and was comfortable with the procedures.

An orientation session was not held for the employees at the beginning of the goal setting treatments. Instead, supervisors introduced the new procedures to each employee individually.

Results

Overall Effects on Performance

Before presenting analyses related to the specific questions addressed by the research, it is instructive to look at overall performance results and to discuss several issues pertinent to their analysis and interpretation. The first issue concerns the unit of analysis. The desired unit of analysis should be long enough to allow for some stability in the performance data; for example, daily performance seemed too small a unit for stability. On the other hand, the amount of time employees worked under a given treatment varied greatly. For example, some employees were hired or left the unit at a time such that only 1 or 2 weeks of data were available for them, while other employees worked for the entire treatment period. To use a long time period, such as performance under an entire treatment, would weight each person equally, no matter how long they worked under the treatment. In an attempt to balance stability and representativeness, it was decided to use weekly mean performance. Thus, unless otherwise noted, the analyses to follow use weekly mean performance as the unit of analysis.

A second issue in the analysis is which employees to use in assessing the magnitude of the treatment effects. One approach would be to

use all subjects who are in the organization at the time of a treatment. The problem with this approach is that due to the fairly high turnover, some employees would be learning the task and learning effects could make interpretation of treatment effects difficult. A second approach is to use only those employees who were present in the units for the baseline period and both treatments. This avoids the problem of the new employees and could give a better estimate of the effects of the treatments. However, to the extent that people are still learning the task, such an analysis may produce inflated treatment effects solely due to improved performance from learning effects. Consequently, it was decided to analyze the data both ways. That is, one set of analyses was done with all the employees who were present in a given treatment whether they were present in other treatment periods or not, and a second set of analyses was done only on those employees who were present for all three periods.

These results are presented in Tables 1 and 2. Table 1 shows the results for Remittance Control for the two types of analyses by shift and condition for each of the three performance variables. A 1 x 3 ANOVA was done on each set of means. That is, the F and p reported in the table are based on the baseline, treatment 1 and treatment 2 means for each performance variable. All these analyses are between-groups analyses. In both types of analyses, but especially in the analyses done on subjects present in all three conditions, there are many cases where observations come from the same subject. However, since the unit of analysis is weekly performance, an equal number of data points was not available for each subject. Thus, the repeated measures aspect of the data set could not readily be used. Consequently, a more conservative between-groups analysis was used.

While the results of the analyses are presented in Tables 1 and 2, they are difficult to interpret due to a problem outlined previously, that is there was a continual flow of individuals in and out of the experimental groups due to new hires and turnover. There was also evidence that learning effects lasting several months were operating on the jobs. Thus, these learning effects could confound the interpretation of the treatment effects. For example, if there were a large number of new hires in the baseline for one group, this could artificially lower the baseline mean for that group. This would make either of the two types of analyses difficult to interpret.

The turnover of the experienced employees and the addition of new employees would not be a problem if the rate of termination and hiring new employees were constant across the three conditions. Unfortunately, this was not the case. Thus, it was necessary to control for learning effects to be able to interpret the treatment effects.

Controlling for learning involved a fairly elaborate process. The first step was to identify the learning curve for each major performance

Table 1

Performance Results: Remittance Control

Analysis	Shift	Variable	Results									
			Baseline		Treatment 1			Treatment 2			F	p
			$\bar{X}$	N*	$\bar{X}$	N	% change from baseline	$\bar{X}$	N	% change from baseline		
All Subjects	Day	Encoding Rate	966.9	387	999.4	656	3.4	995.7	272	3.0	1.9	.15
		Encoding Error Rate	1.93	288	1.95	472	1.0	1.65	185	-14.5	1.8	.16
		Sorting Error Rate	.65	435	.75	746	15.4	.61	296	-6.1	3.9	.02
	Evening	Encoding Rate	1027.3	226	1065.8	369	3.7	1068.3	134	4.0	1.8	.16
		Encoding Error Rate	2.62	160	2.01	245	-23.3	2.1	76	-19.9	3.8	.02
		Sorting Error Rate	.75	236	.66	414	-12.0	.54	134	-28.0	2.7	.07
Subjects Present In All Three Conditions	Day	Encoding Rate	979.5	302	1061.9	467	8.4	1083.9	174	10.6	13.7	.00
		Encoding Error Rate	1.85	212	1.82	313	-1.6	1.46	115	-21.1	2.05	.13
		Sorting Error Rate	.62	332	.70	507	12.9	.57	176	-8.1	1.96	.14
	Evening	Encoding Rate	1013.5	175	1108.6	272	9.4	1102.1	98	8.7	10.0	.00
		Encoding Error Rate	2.73	119	1.90	177	-30.4	2.04	53	-25.3	4.73	.01
		Sorting Error Rate	.75	183	.62	311	-17.3	.52	99	-30.6	2.47	.09

N = The number of weekly performance means, not the number of individuals.

Table 2

Performance Results: Data Input

Analysis	Shift	Variable	Results								
			Baseline		Treatment 1			Treatment 2			F	p
			$\bar{X}$	N*	$\bar{X}$	N	% change from baseline	$\bar{X}$	N	% change from baseline		
All Subjects	Day	Keying Effectiveness	91	548	100	652	9.0	102	268	11.0	46.9	.00
	Evening	Keying Effectiveness	99	375	102	535	3.0	108	222	9.0	15.69	.000
Subjects Present In All Three Conditions	Day	Keying Effectiveness	91	413	100	570	9.0	102	261	11.0	38.9	.000
	Evening	Keying Effectiveness	98	215	104	405	6.0	108	142	10.0	15.21	.000

N = The number of weekly performance means, not the number of individuals.

variable on each job. However, since the feedback and goal setting treatments could influence the shape of the learning curve, it was necessary to develop the learning curve from baseline data. This created a problem in that the baseline was not long enough to follow the same set of people for a long period of time. This problem was overcome by constructing learning curves built up from the data of different people at different points in time. More specifically, a curve was developed showing mean performance by week of tenure on the job. Thus, for example, the point plotted on the learning curve for 9 weeks of tenure consisted of all the people who worked on the job at their ninth week of tenure during the baseline period. The point plotted for the thirtieth week of tenure was based on all those people who worked their thirtieth week of tenure during the baseline, even though this was a different group of people than those making up the mean for the ninth week. Thus, the learning curve as discussed here is not a true learning curve in the sense of the increase in performance over time for a single individual. It does, however, represent expected performance for different amounts of tenure when no treatment is in effect.

The resulting learning curves showed the expected negatively accelerated shape, but were obviously not smooth curves. Because of this, smoothed curves were interposed through the actual data curves. Figure 2 gives an example of the actual data and smoothed curve for keying effectiveness in Data Input. For this variable, performance increases rapidly for the first 5 weeks, increases more slowly for the next 10 weeks, and asymptotes at about 15 weeks. Analogous curves were developed for the three performance variables in Remittance Control.

The next step was to develop the actual corrections for learning. To do this, the asymptotic performance level was calculated. In the case of Data Input, this was the mean keying effectiveness from week 16 through week 52 on the learning curve. Next, each week of tenure on the smoothed learning curve was expressed as a percentage of this asymptotic value. For example, on a given performance variable, performance during the tenth week of tenure might be 83% of asymptotic performance. To make the actual correction, an employee's performance at a given week of tenure was multiplied by a value that would adjust that performance to the asymptotic value. In the above example, if there were data for an employee's tenth week of tenure, the adjusted performance value would be the actual mean weekly performance multiplied by $\frac{100}{83}$. If performance data occurred after asymptote (in Data Input, after 15 weeks) no adjustment was made.

Thus, this procedure adjusts each weekly performance mean to the value to be expected if that person had reached asymptote. It is important to note, however, that since the learning curves were based solely on baseline data, this correction reflects the asymptotic performance that would be expected in the absence of treatment effects. To the

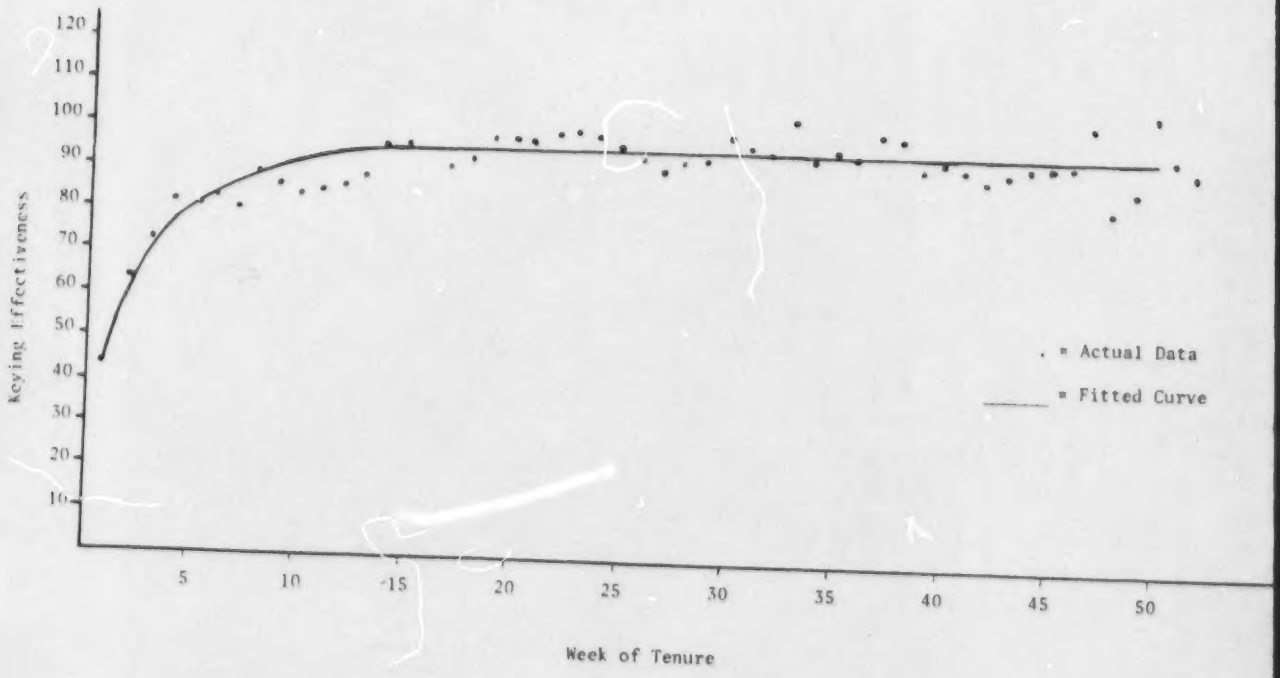


Figure 2 Keying Effectiveness by Week of Tenure in Baseline.

Table 3

Adjusted Performance Results: Remittance Control

Analysis	Shift	Variable	Results									
			Baseline		Treatment 1			Treatment 2			F	p
			$\bar{X}$	N*	$\bar{X}$	N*	% change from baseline	$\bar{X}$	N*	% change from baseline		
All Subjects	Day	Encoding Rate	1009.8	364	1054.4	649	4.41	1030.8	270	2.1	3.4	.03
		Encoding Error Rate	1.82	271	1.88	470	3.3	1.61	184	-11.5	1.6	.20
		Sorting Error Rate	.58	404	.68	726	17.2	.56	294	-3.4	3.4	.03
	Evening	Encoding Rate	1036.5	216	1102.06	363	6.3	1088.8	132	5.0	4.5	.01
		Encoding Error Rate	2.5	148	1.99	240	-20.4	2.09	76	-16.4	2.7	.07
		Sorting Error Rate	.73	224	.61	406	-16.4	.52	133	-28.8	2.8	.06
Subjects Present In All Three Conditions	Day	Encoding Rate	1017.1	301	1041.3	467	6.3	1085.0	174	6.7	7.49	.001
		Encoding Error Rate	1.82	212	1.82	313	0.0	1.46	115	-19.8	1.90	.15
		Sorting Error Rate	.59	331	.69	507	16.9	.57	176	-3.4	2.23	.11
	Evening	Encoding Rate	1035.8	175	1114.1	272	7.5	1102.7	98	6.4	6.54	.002
		Encoding Error Rate	2.68	119	1.9	177	-29.1	2.04	53	-23.7	4.29	.014
		Sorting Error Rate	.71	183	.61	311	-14.1	.52	99	-26.7	1.75	.17

N = The number of weekly performance means, not the number of individuals.

Table 4

Adjusted Performance Results: Data Input

Analysis	Shift	Variable	Results								
			Baseline		Treatment 1			Treatment 2			F	p
			$\bar{X}$	N*	$\bar{X}$	N*	% change from baseline	$\bar{X}$	N*	% change from baseline		
All Subjects	Day	Keying Effectiveness	91	522	99	629	8.0	99	257	8.0	32.70	.000
	Evening	Keying Effectiveness	101	375	105	553	4.0	111	222	10.0	20.81	.000
Subjects Present In All Three Conditions	Day	Keying Effectiveness	91	391	99	547	8.0	99	250	8.0	32.89	.000
	Evening	Keying Effectiveness	100	215	105	405	5.0	108	142	8.0	9.85	.000

N = The number of weekly performance means, not the number of individuals.

Table 5

Treatment Effects: Remittance Control

	Performance Variable	First Treatment			Second Treatment		
		Personal, Comparative Feedback			Personal, Comparative Feedback & Goal Setting		
		All Subjects	Subjects Present In All Three Conditions	Mean	All Subjects	Subjects Present In All Three Conditions	Mean
Day Shift	Encoding Rate	+4.41	+6.31	+5.36	+2.08	+6.68	+3.38
	Encoding Error Rate	+3.39	0	+1.70	-11.53	-19.78	-15.65
	Sorting Error Rate	+17.24	+16.90	+17.07	-3.45	-3.39	-3.42
		Personal, Absolute Feedback			Personal, Absolute Feedback & Goal Setting		
		All Subjects	Subjects Present In All Three Conditions	Mean	All Subjects	Subjects Present In All Three Conditions	Mean
Evening Shift	Encoding Rate	+6.32	+7.56	+6.94	+5.04	+6.41	+5.73
	Encoding Error Rate	-20.40	-29.10	-24.75	-16.40	-23.88	-20.14
	Sorting Error Rate	-16.44	-14.08	-15.26	-28.77	-26.76	-27.76

Table 6

Treatment Effects: Data Input

Performance Variable		First Treatment			Second Treatment		
		Impersonal, Comparative Feedback			Personal, Comparative Feedback		
		All Subjects	Subjects Present In All Three Conditions	Mean	All Subjects	Subjects Present In All Three Conditions	Mean
Day Shift	Format Effectiveness	+8%	+8%	+8%	+8%	+8%	+8%

	Baseline	Personal, Comparative Feedback			Personal, Comparative Feedback & Goal Setting		
		All Subjects	Subjects Present In All Three Conditions	Mean	All Subjects	Subjects Present In All Three Conditions	Mean
Evening Shift	Format Effectiveness	+4%	+5%	+4.5%	+10%	+8%	+9%

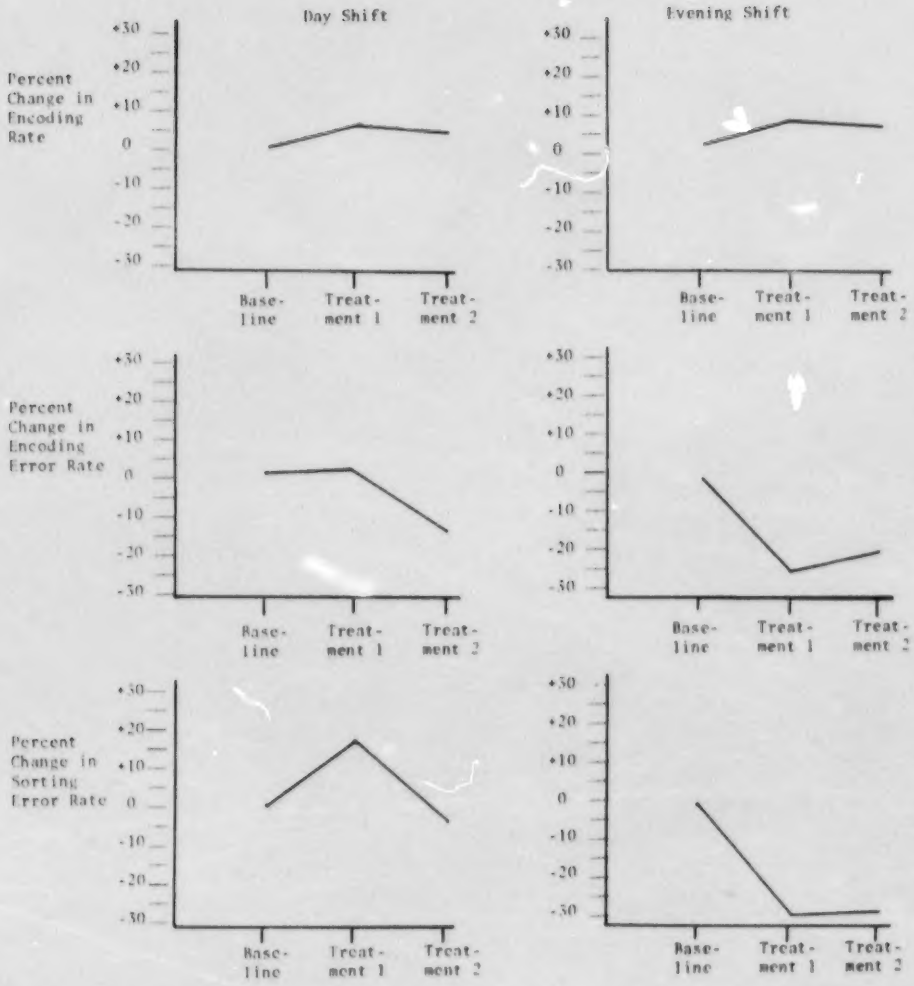


Figure 3 Mean Treatment Effects: Remittance Control.

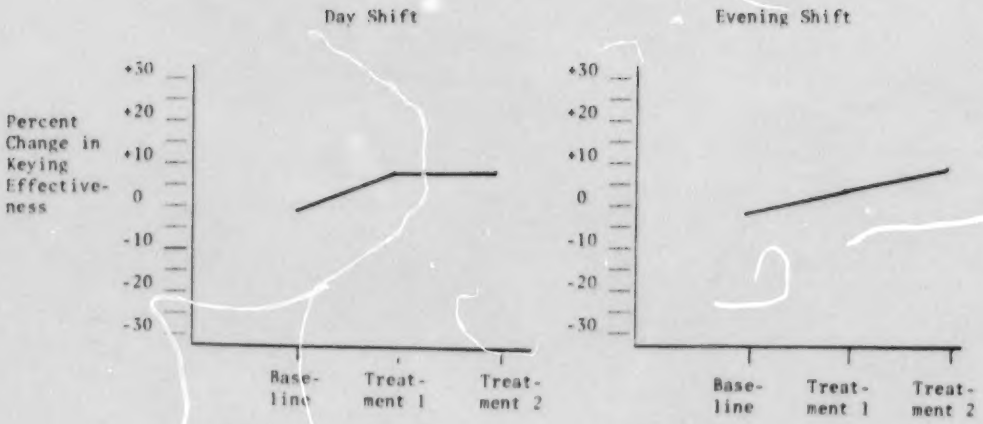


Figure 4 Mean Treatment Effects: Data Input.

extent that adjusted performance under the treatments is higher than adjusted performance under baseline, it can be more safely concluded that true treatment effects occurred, and these effects should be free of bias due to learning effects.

The results of the adjusted performance analyses are presented in Tables 3 and 4. Comparison of these adjusted results with the unadjusted results shows some differences, but not particularly dramatic changes. The general trend is that the treatments had a positive effect on performance and, in some cases, a fairly strong positive effect.

It is also interesting to compare the results of the two types of analyses after the learning curve correction. Tables 5 and 6 display the percent change from baseline performance for the analysis using all subjects present during each of the conditions and the analysis using only subjects present in all three conditions. Comparison of these two types of analyses shows rather good convergence. The effect sizes are quite similar for the two types of analyses. This lends confidence to our assessment of the magnitude of the treatment effects. Also presented in these tables are the mean effect sizes for the two types of analyses. These means reflect the best estimate of the magnitude of the treatment effects and are plotted in Figures 3 and 4.

These results indicate that the treatments generally showed a positive effect on performance. In Remittance Control, 10 of the 12 changes from baseline were in the direction of better performance. Performance did not improve in the error data in the personal, comparative feedback condition on the day shift. However, this increase in error rate was possibly an artifact due to circumstances on the job at that time. Specifically, there was a particularly heavy inflow of work during the last two-thirds of this treatment. As a result of this increased workload an unusually high number of temporary employees were hired. In addition, the usual procedure of employees rotating back and forth through the day from the encoding task to the sorting task was suspended, and employees typically worked on only one task or the other for the entire day. In addition, the amount of overtime increased substantially. One consequence of this might be increased fatigue and a consequent increase in errors. Several factors support the interpretation that the increased workload was responsible for the increase in error rates. First, it was only after this increased workload occurred that error rates increased. Second, a similar treatment (personal absolute rather than personal comparative feedback) in the evening shift was not affected by the increased workload and showed no decrement. Third, the same treatment when used in Data Input showed positive effects. Thus, it is concluded that the increase in error rates in the personal comparative feedback treatment in Remittance Control was not due to the treatment, but rather to the increased workload. The results for Data Input show consistent positive effects on performance for the treatments. In every case, the treatments showed increased performance over the baseline.

A second measure of performance in Data Input was a type of error rate measure. However, in contrast to the error rate measures in Remittance Control, the measure in Data Input was a group measure that could not be tied to specific individuals. Nevertheless, it is instructive to examine this index to see if there is any evidence that the increase in quantity of output in Data Input was accompanied by an increase in errors.

The results suggest that this was not the case. The index of errors per unit of work in the day shift (or baseline, treatment 1, and treatment 2 was .13, .13, and .11, respectively. Analogous values for the evening shift were .20, .18, and .20. Thus, if anything, there was a trend toward decreasing error rates during the treatments.

Evaluations of Specific Treatments.

Previous sections have considered the overall performance effects, and now the specific treatments will be evaluated. These are personal versus impersonal feedback, absolute versus comparative feedback, and goal setting.

Three comparisons are possible in the evaluation of personal versus impersonal feedback. The best test is the two treatments for the day shift in Data Input. Here there is no difference in performance whatsoever. Also, of course, the a priori contrast shows no difference. (The mean contrasts that follow, will all be a priori and use all subjects in the treatment at that time. Results for the analyses for subjects present in all conditions were essentially identical. All contrasts are done on the adjusted performance data.) The second comparison is in the first treatment in Data Input, comparing performance in the day shift (impersonal) to performance in the evening shift (personal). The contrast here is the difference between baseline and treatment for one shift compared to the difference for the other shift. This comparison shows impersonal feedback significantly better ($F = 31.77$, $p < .001$) than personal. The third comparison is between the first treatment, evening shift of Remittance Control and the first treatment, day shift of Data Input. Here we are comparing across jobs with different dependent variables, so a significance test is not appropriate. However, if we look only at quantity of output, personal feedback showed a greater increase over baseline than did impersonal. Taking these three comparisons together, the results suggest that the personal and impersonal feedback do not result in different levels of performance.

The second treatment comparison is between absolute and comparative feedback. The first contrast is between the day and evening shifts of the first treatment in Remittance Control. Here again the change from baseline for the two treatments is being compared. For encoding rate, there was no significant difference ($F = .55$, n.s.); there were significant differences in favor of absolute feedback for encoding error rate

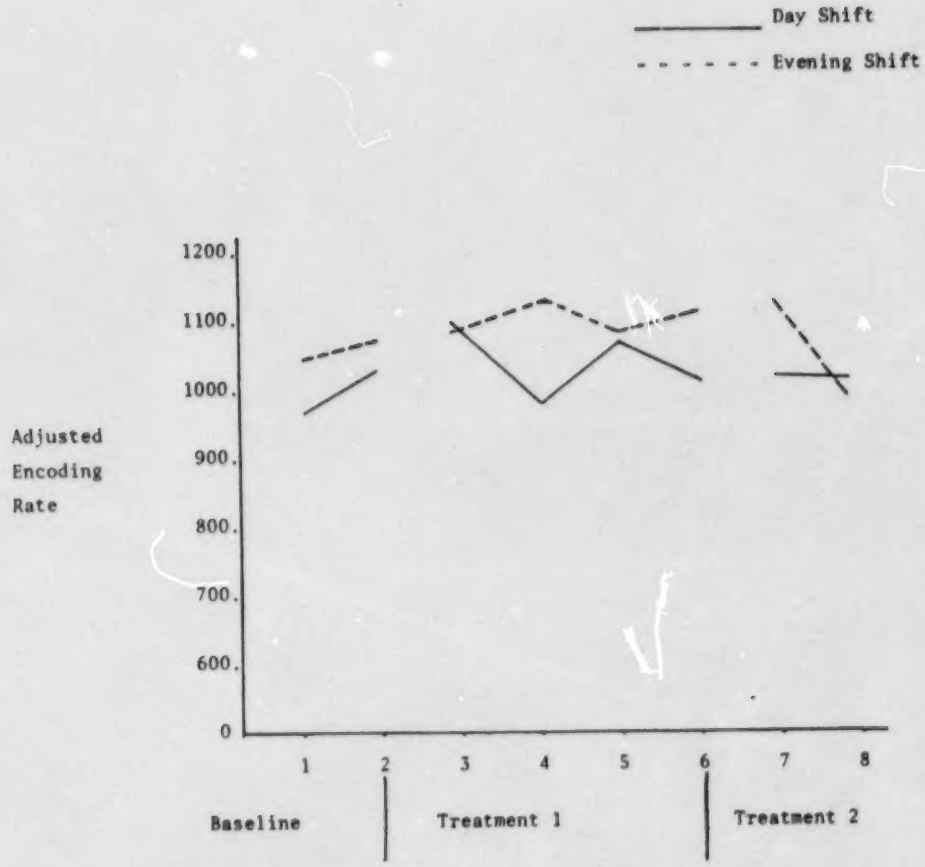


Figure 5 Performance by Month: Remittance Control, Adjusted Encoding Rate.

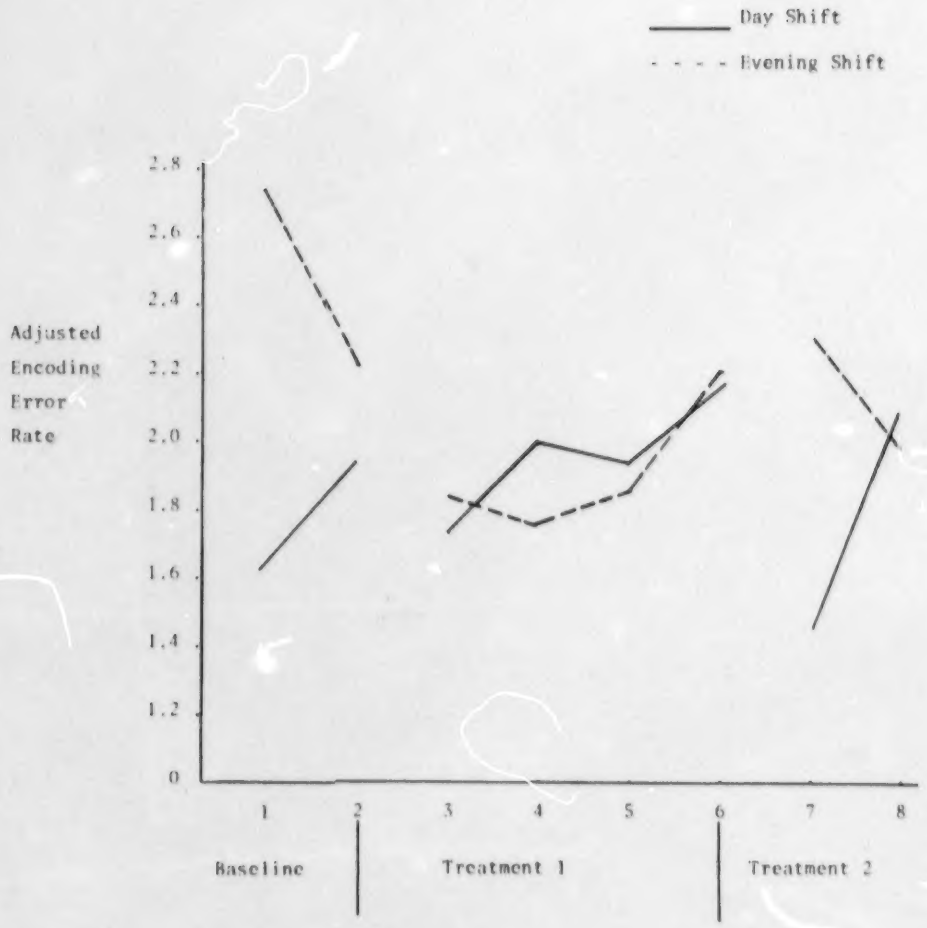


Figure 6 Performance by Month: Remittance Control, Adjusted Encoding Error Rate.

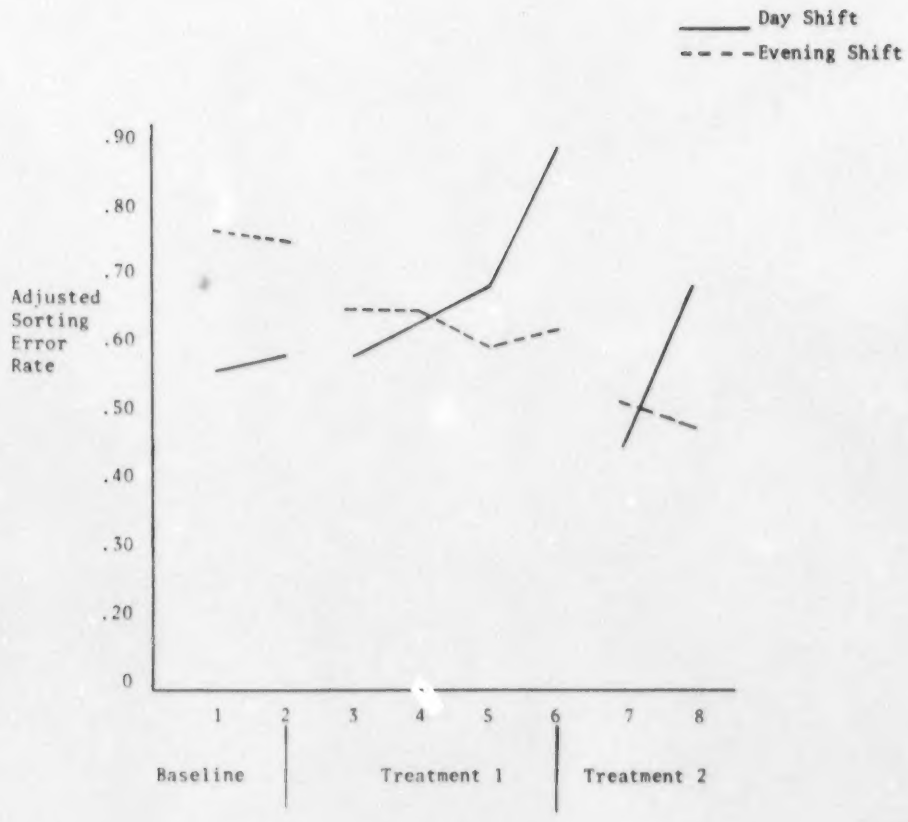
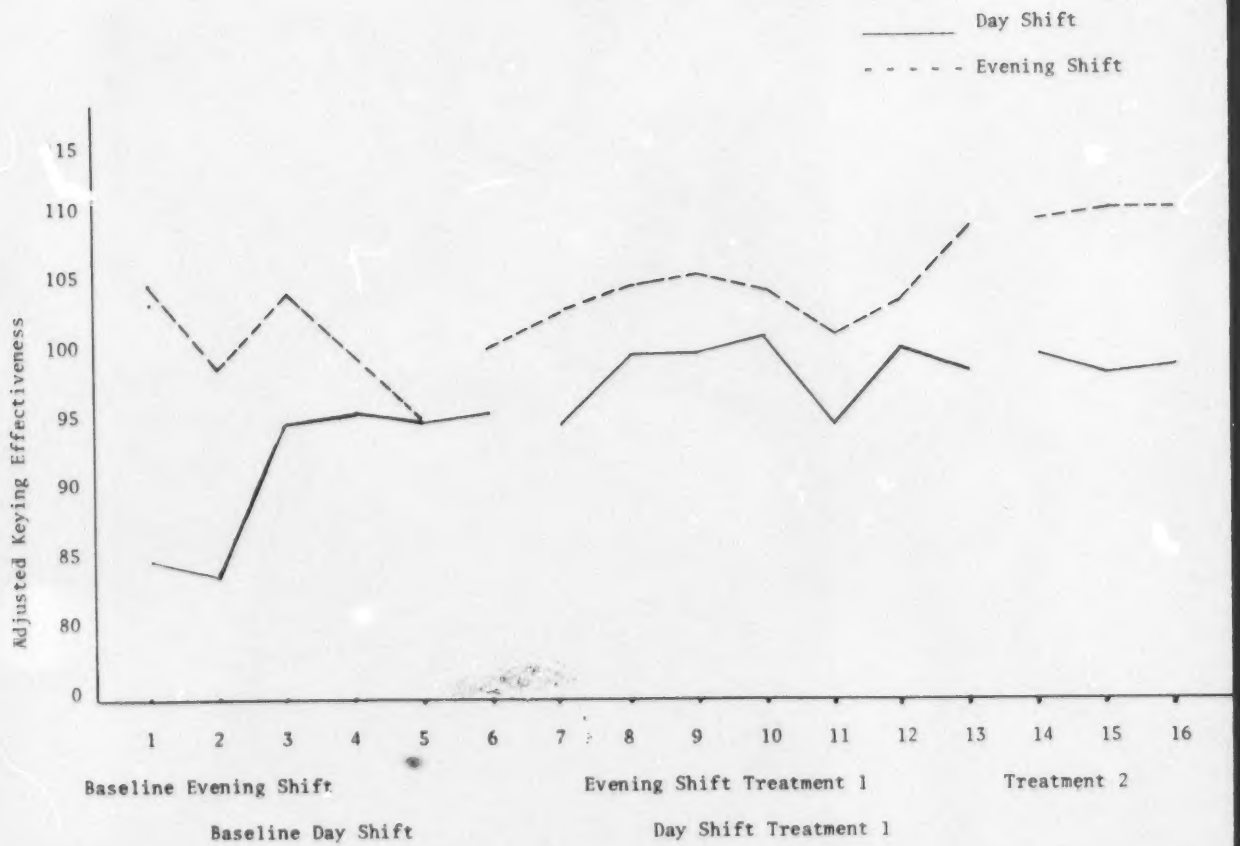


Figure 7 Performance by Month: Remittance Control, Adjusted Sorting Error Rate.

Figure 8 : Performance by Month: Data Input
Adjusted Keying Effectiveness



($F = 5.75$, $p < .05$) and for sorting error rate ($F = 8.64$, $p < .01$). However, before placing too much confidence in these results, it must be remembered that the comparative condition is where the work overload occurred. As has been argued previously, this overload probably increased error rates artificially. Two other across-job comparisons are possible for absolute and comparative feedback. The first is between the first treatment, evening shift of Remittance Control and the first treatment, evening shift of Data Input. Here there is some superiority of absolute over comparative in terms of quantity of output. The second across-job comparison is between the same condition in Remittance Control and the second treatment in the day shift of Data Input. Here, the results favor comparative feedback. Taken together, and bearing in mind the problems with the error rates in day shift of the first treatment in Remittance Control, the results suggest that neither absolute nor comparative feedback is superior.

The third type of treatment is goal setting. Here, there are three within-job comparisons where goal setting was added to a feedback condition. The first is the day shift in Remittance Control. Comparing feedback to feedback plus goal setting showed no significant difference in encoding rate ($F = 1.50$, n.s.), no significant difference for encoding error rate ($F = 3.04$, n.s.), and a significant difference in favor of goal setting for sorting error rate ($F = 5.14$, $p < .05$). The second comparison is with the evening shift of Remittance Control. Here, none of the three variables were significantly different ($F = .26$, .13, and 1.35, respectively). The final comparison is in the evening shift of Data Input. Here, goal setting showed significant improvement over feedback ($F = 288.46$, $p < .001$).

Taken as a whole, the results support the notion that goal setting improved performance over feedback. While most of the differences were not significant, those that were did support goal setting, and the remainder were generally in the predicted direction.

Effects of Treatments Over Time

The next issue to be explored in the results is the effects of the treatments over time. The issue here is whether the treatments tended to show diminishing results over time. One possibility is that the treatments had an initial positive effect on performance, but after the novelty was gone, performance decreased. To address this question, each treatment period was broken down into 4-week periods, and performance means were calculated for each period within each treatment.

Figures 5, 6, 7, and 8 present the results of these analyses. The plotted means represent the mean adjusted performance by four-week periods for all subjects who were present at the time. Results for employees present in all three conditions were essentially similar. The

Figure 9 : Treatment Effects by Initial Level of Performance: Remittance Control, Encoding Rate

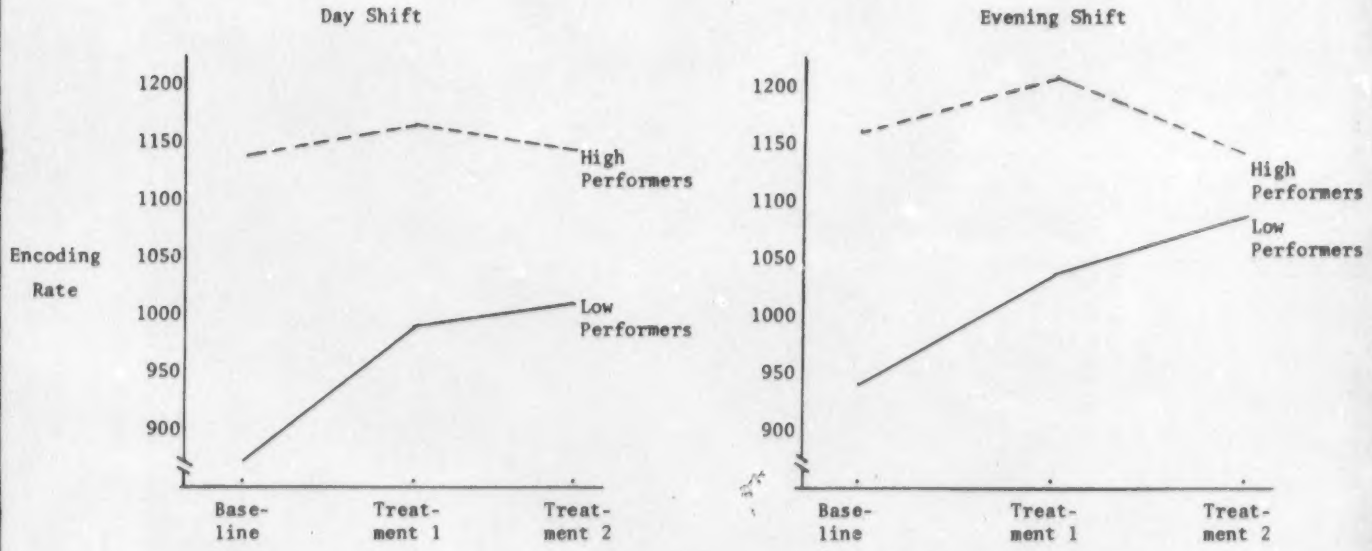


Figure 10: Treatment Effects by Initial Level of Performance: Remittance Control, Encoding Error Rate

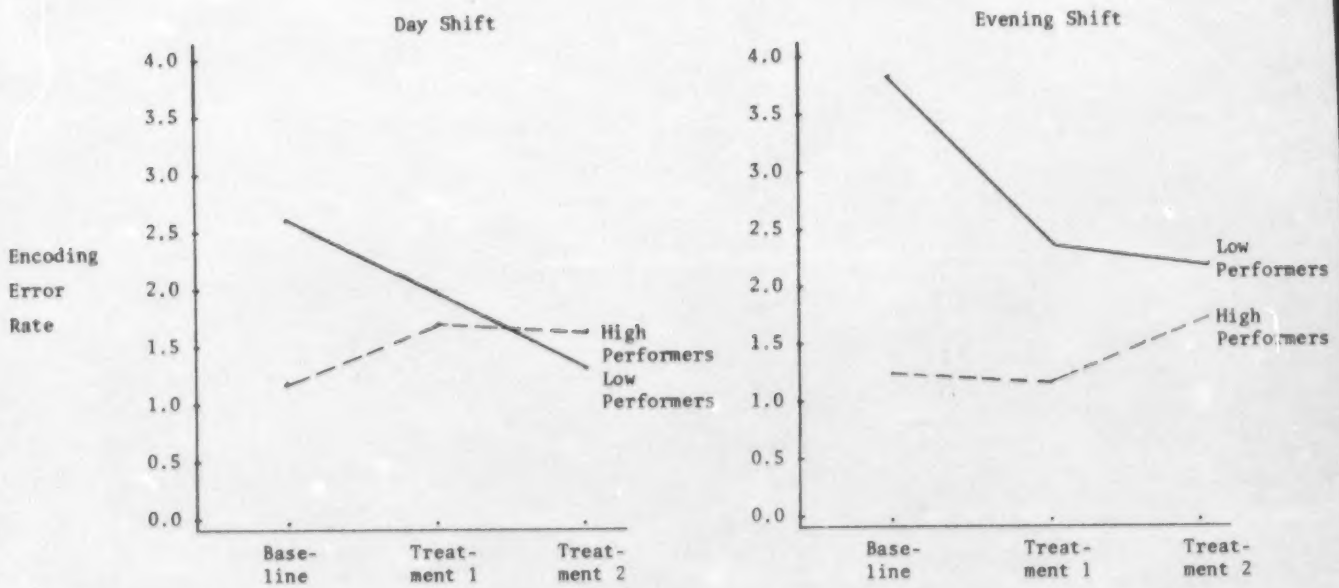


Figure 11: Treatment Effects by Initial Level of Performance: Remittance Control, Sorting Error Rate

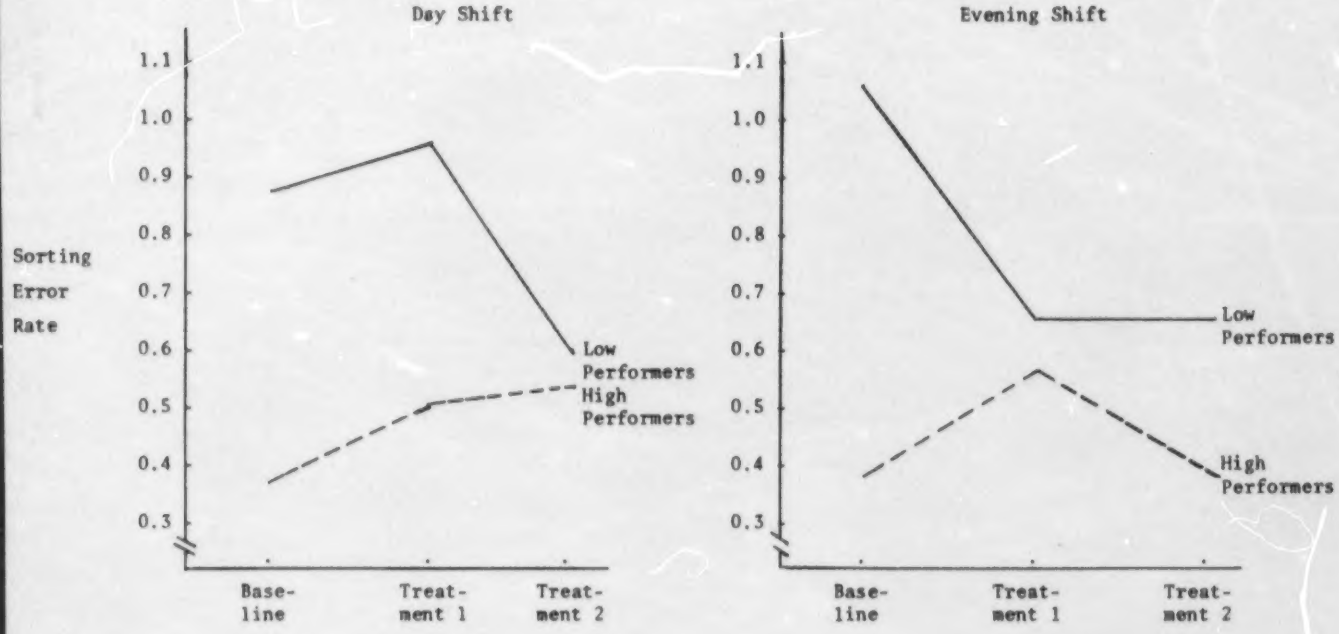


Figure 12: Treatment Effects by Initial Level of Performance: Data Input, Keying Effectiveness

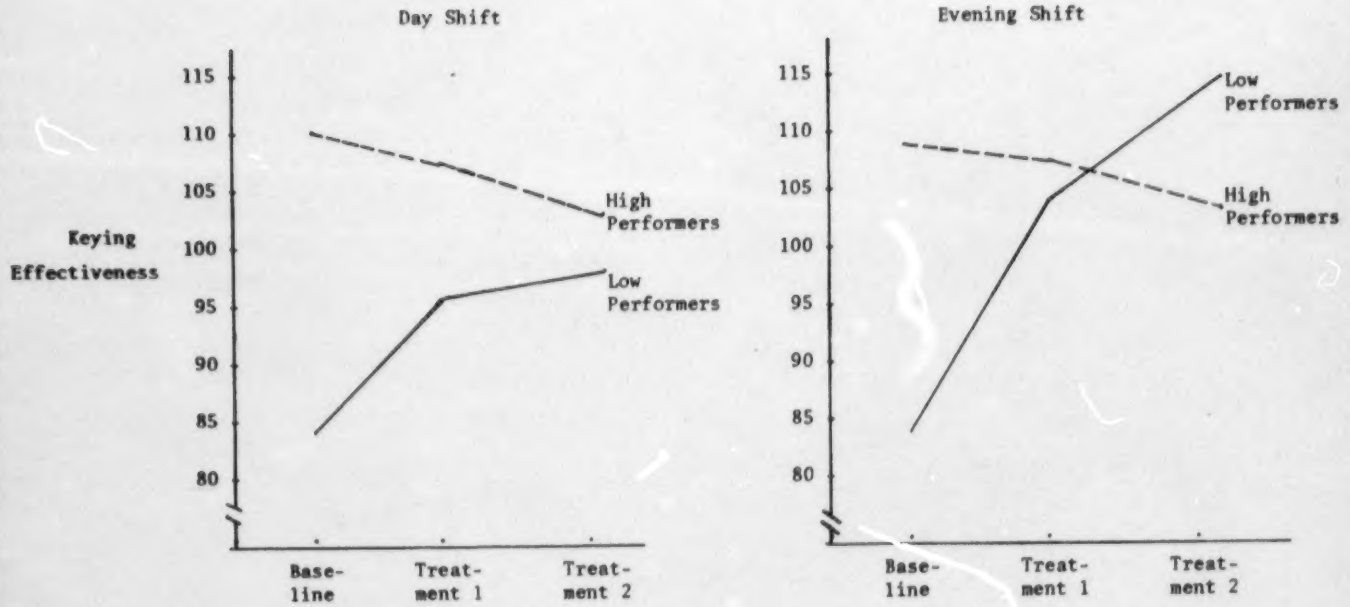


Figure 13: Performance by Tenure: Remittance Control, Encoding Rate.

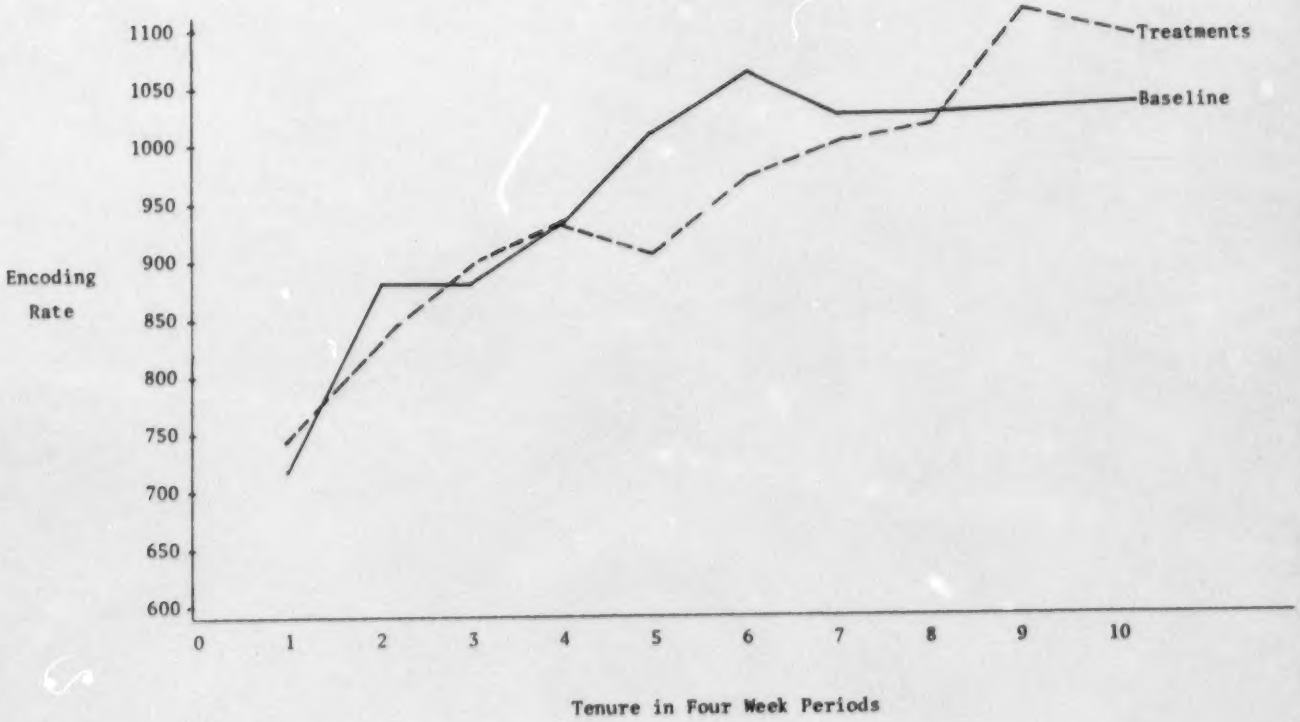


Figure 14 : Performance by Tenure: Remittance Control, Encoding Error Rate

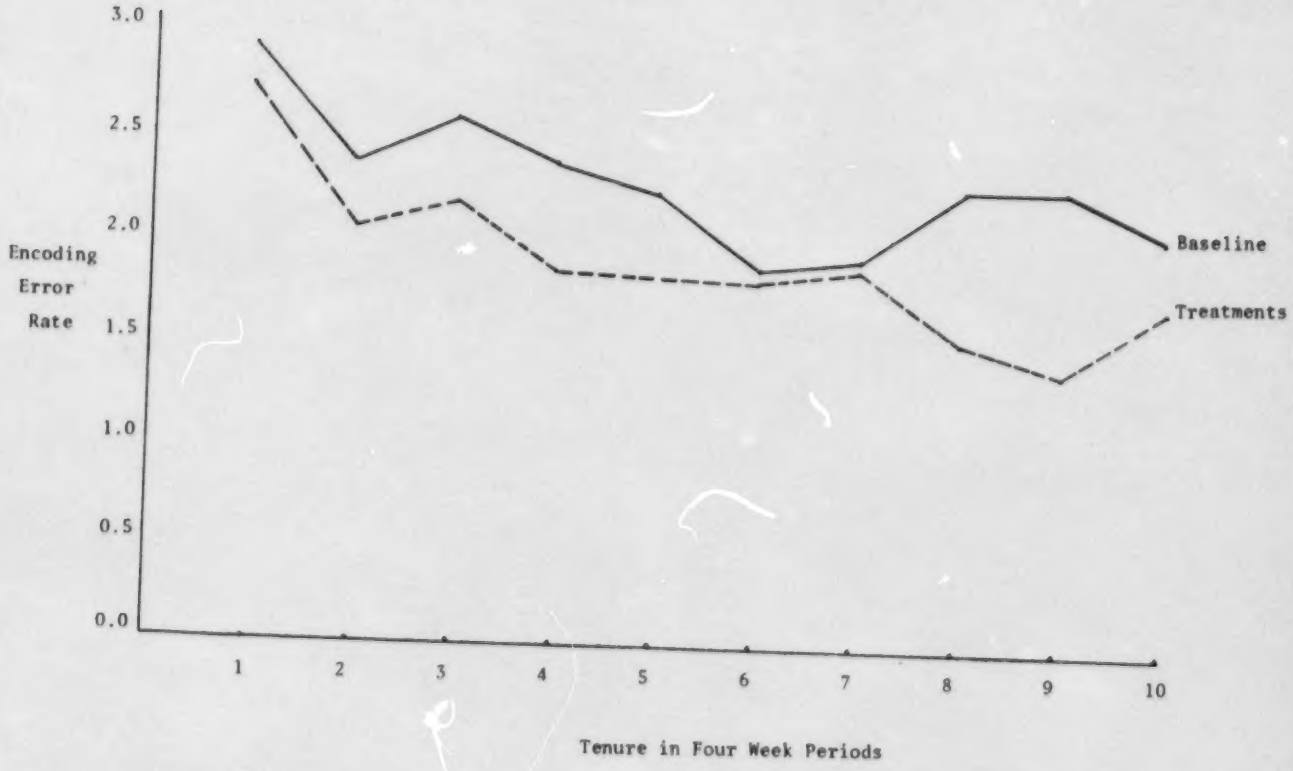


Figure 15: Performance by Tenure: Remittance Control, Sorting Error Rate

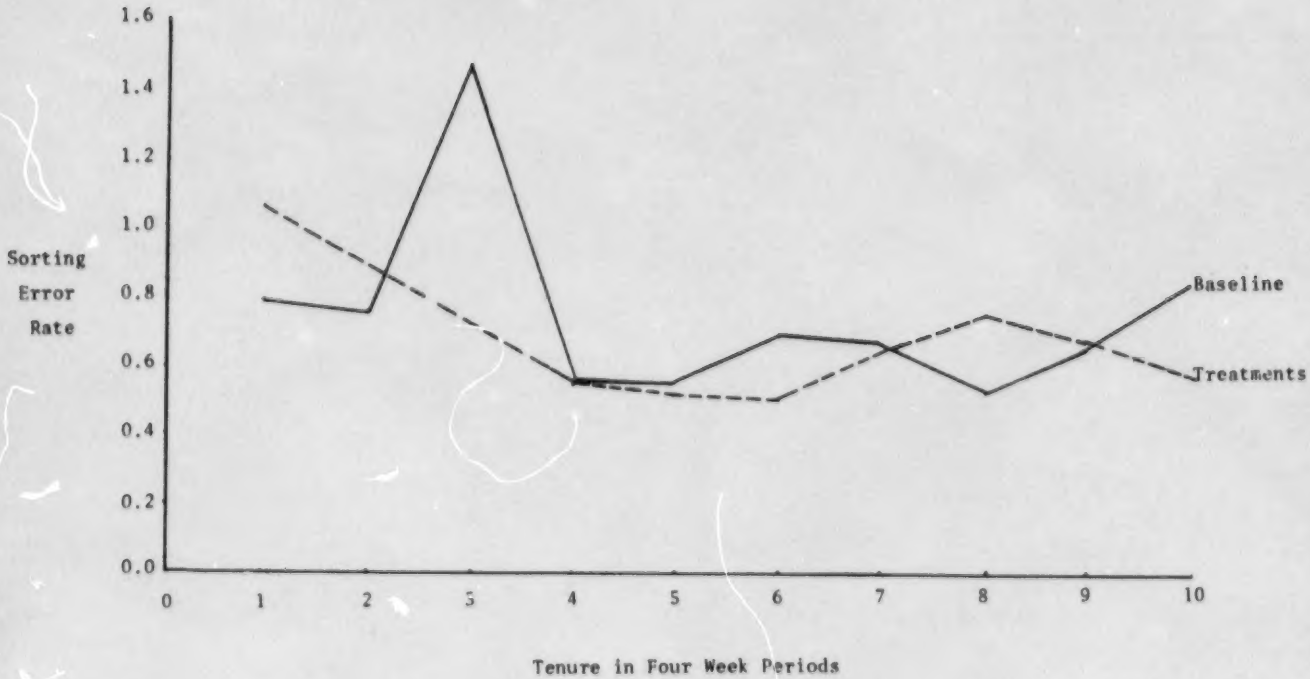
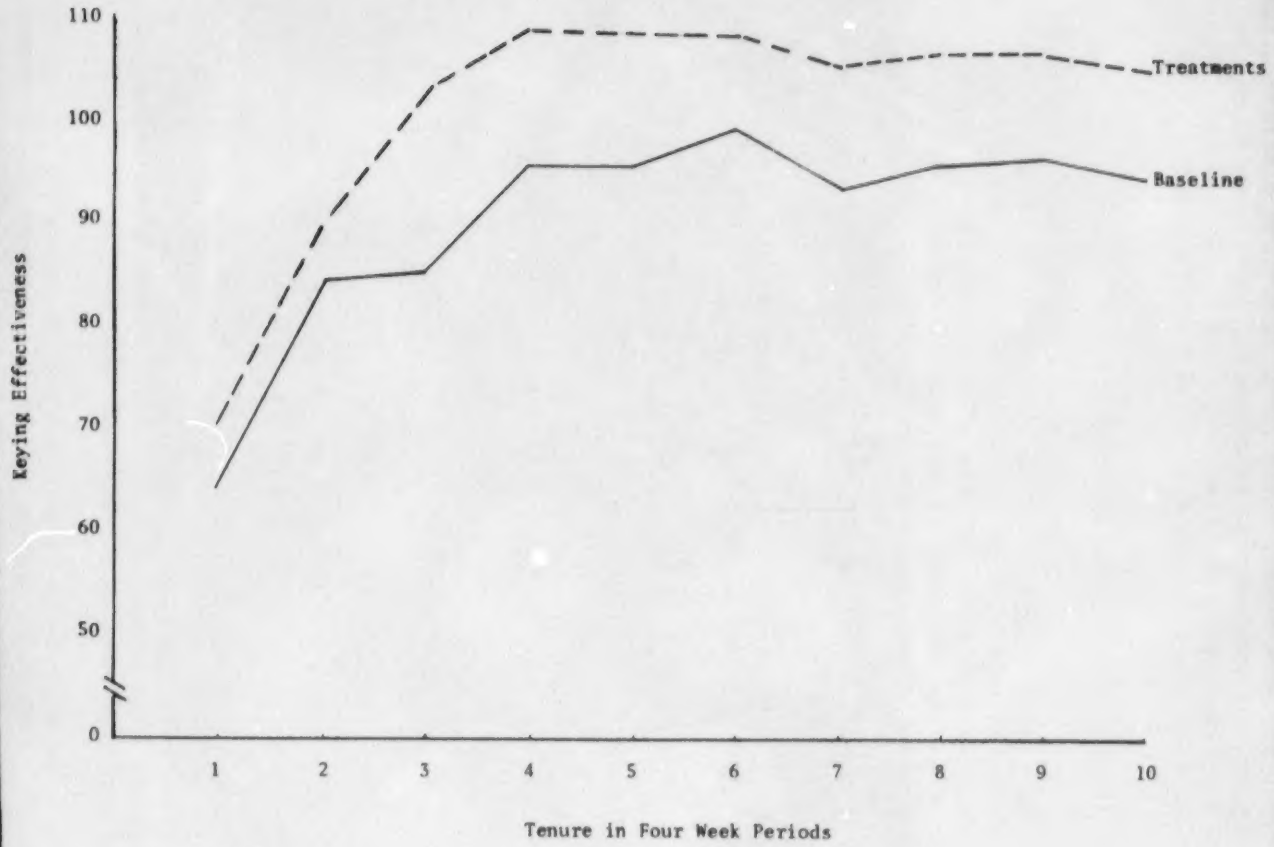


Figure 16 : Performance by Tenure: Data Input



first three figures present the results for Remittance Control. Inspection of these means shows considerable variability within each treatment condition, but there is no general pattern of decreasing performance within the treatments. The results for Data Input are presented in Figure 8. Here again, there is no pattern of decreasing performance through the treatments. If anything, there is a trend for performance to increase during the treatments; thus, the evidence suggests that the effects of the treatments were not diminishing over time.

Effects of Ability on Performance

Another question of interest was the effects of the treatments as a function of ability. That is, whether the treatments had differential effects on employees of different ability. Unfortunately, adequate independent ability measures were not available. Consequently, performance during the baseline was used to classify employees within each shift as high or low depending on whether they were above or below median performance. The performance of these two groups for the two experimental conditions was then calculated. Because performance was used to classify people, it is not totally appropriate to talk about high and low ability. The two groups could just as well be classified as high and low motivation groups. Consequently, the designation high and low performing groups was used.

Figures 9, 10, 11, and 12 present the results of these analyses. Inspection of these figures shows a clear and consistent pattern. The treatments had little positive effect on subjects who were initially high performers. Almost all the effects come from those subjects who were initially low performers. Thus, the treatments increased the performance of the low performers, but had little effect on the high performers.

Effects on Learning

The final aspect of the performance results to be considered is the effects of the feedback and goal setting treatments on learning. The issue is whether the treatments have any effect on the speed of learning the task. To explore this issue, learning curves were built up in the same fashion as described earlier. That is, each person's performance during the treatments was coded as to the number of weeks of tenure on the task each weekly performance mean represented. The mean of these values was then taken as the data point for each week of tenure. As before, this procedure results in a learning curve that is a composite of different people at different times.

The results of these analyses are presented in Figures 13, 14, 15, and 16. These figures represent the learning curve present in the baseline and a second curve for the treatments. The treatment curve is

collapsed across shift and treatment due to the small sample size. That is, within a single shift and treatment, it was not uncommon to have only a few people at a given number of weeks of tenure. In addition, the curves are plotted in 4-week periods to show the effects more clearly.

Inspection of these figures indicates mixed results. For encoding rate and sorting error rate in Remittance Control, there do not seem to be any effects on the rate of learning. For these variables, the baseline and treatment curves are quite similar. For encoding error rate in Remittance Control and keying effectiveness in Data Input, learning appears to be occurring more rapidly under the treatments. For encoding error rate, the effect is present throughout the learning period. For keying effectiveness, the effect occurs after the third 4-week period.

Attitude Results

Subjects were given the Minnesota Satisfaction Questionnaire (MSQ) during the baseline and each treatment. Results for overall satisfaction are presented in Table 7. In only one of the four groups were there significant effects using a 1 x 3 ANOVA. In this group, Remittance Control, day shift, job satisfaction was higher in the treatments than in the baseline. In the other three groups, there were no significant results. Inspection of the individual items of the MSQ showed no interpretable effects.

Summary of the Results

Overall, the results can be summarized as follows:

1. The treatments showed an overall positive effect on performance. Increases in quantity of output typically ranged from 5% to 10% with a mean increase of 6.4%. Error rates decreased. The mean decrease in errors was 11%, with over half the decreases in the 15% to 28% range.

2. Personal feedback was equally as effective as impersonal feedback.

3. Absolute feedback was equally as effective as comparative feedback.

4. Goal setting plus feedback showed higher performance than feedback alone.

5. The positive effects of the treatments did not diminish over time.

6. The treatments had fairly strong effects on employees who were

Table 7

Mean Overall Job Satisfaction

		Baseline		Treatment 1		Treatment 2		F	P
		$\bar{X}$	N	$\bar{X}$	N	$\bar{X}$	N		
Remittance Control	Day	3.09	60	3.36	55	3.26	47	3.00	.05
	Evening	3.27	31	3.12	30	3.27	22	.41	.67
Data Input	Day	3.14	36	3.29	29	3.08	27		.51
	Evening	3.24	23	3.12	23	3.34	18		.52

initially low performers. They did not have much effect on employees who were initially high performers.

7. There was some evidence that the treatments affected the rate of learning, but these results were not present in all situations.

8. Attitudes under the treatments were as good or better than before the treatments.

Discussion and Conclusions

This field test of the earlier laboratory work has shown that feedback can be an effective means of increasing productivity while attitudes remain as good or better than before feedback. The various conditions showed consistent positive effects on quantity and errors.

More specifically, the results show that personal and impersonal, as well as absolute and comparative feedback, are equally effective. Thus, the choice of which type of feedback to use would be best determined by the particular circumstances and preferences of the unit to which the feedback would be applied.

The best results were obtained when feedback was combined with goal setting. However, the goal setting process requires more management resources. Whether this is a cost effective procedure would depend on the difficulty of utilizing management time to conduct the goal setting sessions. One issue here is the frequency of goal setting sessions. A 2-week period was used in this project. It might be possible to use less frequent goal setting sessions and still find substantial positive effects.

One interesting finding is that the treatments had more positive effects on poorer performers. If it is assumed that higher performers are in general more highly motivated, this finding is not surprising. The treatments are designed to increase motivation. If the high performers are highly motivated to start with, the treatments would have less effect on them. It would be the employees lower in motivation who would be most affected.

These findings would appear to have direct applicability to Air Force settings. The jobs used in this study, while civilian jobs, have direct counterparts in the Air Force. Jobs where there is some combination of repetitive manual processing combined with cognitive processing are quite common in the Air Force. Furthermore, there is no reason to think that the principles of feedback and goal setting would not apply to any type of job where repetition is fairly high.

In addition, feedback and goal setting are fairly straightforward

to apply. Developing good feedback and goal setting systems does not require extensive training, and once the principles are learned, these systems could be implemented by local managers.

There are several such general principles that can be recommended about such feedback and goal setting systems. The first is that feedback should be frequent, objective, and specific. The frequency will be a function of the time it takes to complete a unit of work. If this time, or job cycle, is short, e.g., from a few minutes to a few hours, daily feedback seems preferable. If it is longer, feedback should occur less frequently. The more objective the feedback, the better. Although this is not always possible, objectivity should be a goal. Feedback should also be specific to the individual; that is, it should deal with that individual's own work.

A second principle applies to the scope of the feedback. There should be separate feedback for each distinct type of activity on the job. That is, if the person does qualitatively different kinds of tasks, there should be separate feedback for each one. In addition, however, it would seem advisable to have an overall index of effectiveness as well.

The final general feedback principle is probably the most important. The system should be designed so that individuals can see that changes in their levels of effort result in changes in their levels of performance. This is critical to an effective system. As an example of this principle, in the job used in this study, people received considerable feedback before the treatments. For example, in Data Input, employees were given their average keystrokes per hour each day. However, this index had little meaning or utility for them because the difficulty level of the work varied considerably. By generating an index that took difficulty level into account, the employees were able to better see a connection between their effort and their performance.

Once a good feedback system is designed, adding goal setting to it is fairly straightforward. But based on this and other research, the goal setting should be individual, it should take ability into account, it should be done with the employee participating in the process, it should be specific and quantitative, and it should have the goals set as high as the person feels is reasonable to obtain.

In conclusion, feedback and goal setting systems show clear potential for increasing productivity in Air Force settings. They are also productivity enhancing techniques that are implementable at the local level and, thus, can serve as feasible tools for the individual managers.

References

Ammons, R.B. Effects of knowledge of performance: A survey and tentative theoretical formulation. Journal of General Psychology, 1956, 54, 279-299.

Annett, J. Feedback and human behavior. Baltimore: Penguin, 1969.

Beeson, P. Immediate knowledge of results and test performance. Journal of Educational Research, 1973, 66, 224-226.

Berman, M., Fraser, M.P., & Theious, J. Learning a general maximum likelihood decision strategy. Journal of Experimental Psychology, 1970, 86, 393-397.

Christal, R.E. The United States Air Force occupational research project. AFJRL-TR-73-75.

Christian, W.P.,Jr. Conditioned attraction toward neutral peers in a task oriented situation. Representative Research in Social Psychology, 1972, 3, 121-129.

Deci, E.L. The effects of externally mediated rewards on intrinsic motivation. Journal of Personality and Social Psychology, 1971, 18, 105-115.

Fisher, C.D., & Pritchard, R.D. Effects of personal control, extrinsic rewards, and competence on intrinsic motivation. AFHRL-TR-78-20, AD-A058 417. Brooks AFB, TX: Occupation and Manpower Research Division, Air Force Human Resources Laboratory, July 1978.

Gould, R.B. Review of an Air Force job satisfaction research project: status report through September, 1976. AFHRL-TR-76-76.

Greller, M.M. Feedback and reactions to the job. Report # 75-98. New York: New York University, Graduate School of Administration, 1975.

Greller, M.M., & Herold, P.M. Sources of feedback: A preliminary investigation. Organizational Behavior and Human Performance, 1975 (April), 13(2), 244-256.

Greller, M.M., & Herold, P.M. Feedback: Definition of a construct. Academy of Management Journal, 1977, 20, 142-147.

Griswold, P.M. Feedback and experience effects on psychological reports and predictions of behavior. Journal of Clinical Psychology, 1970, 26, 439-442.

Hackman, J.R., & Lawler, E.E.,III. Employee reactions to job characteristics. Journal of Applied Psychology, 1971, 55, 259-286.

Hackman, J.R., & Oldham, G.R. Development of the job diagnostic survey. Journal of Applied Psychology, 1975, 60, 159-170.

Ivancevich, J.M., Donnelly, J.N., & Lyon, J.L. A study of the impact of management by objectives on perceived need satisfaction. Personnel Psychology, 1970, 23, 139-151.

Ilgen, D.R., Fisher, C.D., & Taylor, M.S. Consequences of individual feedback on behavior. Journal of Applied Psychology, 1979, 64, 349-371.

Lair, C.V., & Smith, R.K. Utilization of verbal and non-verbal feedback among college Ss. Psychological Reports, 1970, 27, 807-810.

Latham, G.P., & Yukl, G.A. A review of research on the application of goal setting in organizations. Academy of Management Journal, 1975, 18, 824-845.

Leskovec, E.W. A guide for discussing the performance appraisal. Personnel Journal, 1967, 46, 150-152.

Locke, E.A. Toward a theory of task motivation and incentives. Organizational Behavior and Human Performance, 1968, 3, 157-189.

Locke, E.A., Cartledge, N., & Koeppel, J. Motivational effects of knowledge of results: A goal-setting phenomenon. Psychological Bulletin, 1968, 70, 474-485.

Longstreth, L.E. Test of the law of effect using open and closed tasks. Journal of Experimental Psychology, 1970, 84, 53-57.

McCall, M.W.,Jr., & DeVries, D.L. When nuts and bolts are not enough: An examination of the contextual factors surrounding performance appraisal. Paper presented at the 84th Annual Convention of American Psychological Association, Washington, D.C., September 1976.

Merrill, D.M. Specific review versus repeated presentation in a programmed imaginary science. Journal of Educational Psychology, 1970, 61, 392-399.

Payne, R.B., & Hauty, G.T. The effect of psychological feedback upon work decrement. Journal of Experimental Psychology, 1955, 50, 343-351.

Porter, L.W., & Lawler, E.E.,III. Managerial attitudes and performance. Homewood, IL: Dorsey Press, 1968.

Pritchard, R.D., DeLeo, P.J., & Von Bergen, C.W. Incentive motivation techniques evaluation in Air Force technical training. AFHRL-TR-74-24, AD-A005 302. Lowry AFB, CO: Technical Training Division, Air Force Human Resources Laboratory, November 1974.

Pritchard, R.D., & Montagno, R.V. The effects of specific vs. non-specific and absolute vs. comparative feedback on performance and satisfaction. AFHRL-TR-78-12, AD-A055 693. Brooks AFB, TX: Occupation and Manpower Research Division, Air Force Human Resources Laboratory, May 1978.

Pritchard, R.D., Montagno, R.V., & Moore, J.R. Enhancing productivity through feedback and job design. AFHRL-TR-78-44, AD-A061 703. Brooks AFB, TX: Occupation and Manpower Research Division, Air Force Human Resources Laboratory, August 1978.

Robinson, J.A., & Kulp, R.A. Knowledge or prior recall. Journal of Verbal Learning and Verbal Recall, 1970, 9, 84-88.

Sassenwrath, J.M., & Younge, G.D. Effects of delayed information feedback cues in learning on delayed retention. Journal of Educational Psychology, 1969, 60, 174-177.

Simpson, W.E. Latency of locating lights and sounds. Journal of Experimental Psychology, 1972, 93, 169-175.

Steers, R.M., & Porter, L.W. The role of task-goal attributes in employee performance. Psychological Bulletin, 1974, 81, 434-452.

Thorndike, E.L. The law of effect. American Journal of Psychology, 1927, 39, 212-222.

Turner, A.N., & Lawrence, P.R Industrial jobs and the worker: An investigation of responses to task attributes. Boston: Graduate School of Business Administration, Harvard University Division of Research, 1965.

Tuttle, T.C., Gould, R.B., & Hazel, J.T. Dimensions of job satisfaction: initial development of the Air Force Occupational Attitude Inventory. AFHRL-TR-75-1.

Weidner, F. Personal and impersonal social reinforcement. Psychological Reports, 1968, 23, 885-886.

Weiss, D.J., Dawis, R.V., England, G.W., & Lofquist, L.H. Manual for the Minnesota Satisfaction Questionnaire. Minnesota Studies in Vocational Rehabilitation, 1967, Vol. XXII.

Wike, S.S. The effects of feedback, guessing and anticipation rate upon verbal discrimination learning. The Psychological Record, 1970, 20, 171-178.

Appendix A:

Example Feedback Report for Remittance Control

REMITTANCE CONTROL OPERATOR STATISTICS REPORT

OPERATOR NUMBER 75.

DAILY RAW INFORMATION

SORTING DATA

TIME SORTING (MINUTES) 0.

ENCODING DATA

6061 ENCODES

TOTAL 6061 ENCODES	2199.
TIME ENCODING 6061 (MIN)	195.
606[illegible] ENCODE RATE (PER HR)	677.
ERROR RATE (PER 100 ENCODES)	4.65

638TATEMENT ENCODES

TOTAL 638TMT ENCODES	0.
TIME ENCODING 638TMT (MIN)	0.
638T ENCODE RATE (PER HR)	0.
ERROR RATE (PER 100 ENCODES)	*****

63CHECK ENCODES

TOTAL 63CHK ENCODES	3637.
TIME ENCODING 63CHK (MIN)	205.
63CH ENCODE RATE (PER HR)	1064.
ERROR RATE (PER 100 ENCODES)	0.79

OPEX DATA

TOTAL OPEX SORTS	0.
TIME SORTING OPEX (MI[illegible]	0.
OPEX SORTING RATE (PER [illegible])	0.

OPERATOR DAILY FEEDBACK

	DAY OF 103180	PREVIOUS WEEK
SORTING ACTIVITY		
ERROR RATE (PER 100 SORTS)	*****	******
OPEX SORTING RATE	0.	0.
ENCODING ACTIVITY		
AVERAGE ENCODE RATE (ENCODES/HR)	997.	880.
AVERAGE ERROR RATE (ERRORS/100 ENCODES)	2.92	1.85
OVERALL EFFECTIVENESS INDEX	77.	73.

Appendix B:

Example Feedback Report for Data Input

A-SHIFT 06-02-80
OPR 054 H THOMAS

KEY-MIN	%-OF-MINIMUM	RANK
284	72%	19 OF 24

SEARCH-MIN	RANK
*	12 OF 24

EXCEPTION-TIME
20

UNACCOUNTED-FOR-TIME	RANK
155	3 OF 24

FORMAT-EFFECTIVENESS	RANK
94%	15 OF 24

FORMAT-PERFORMANCE

FORMAT	KS-HR	%-OF-STD	V	TIME H	M
C-001	11,070	91%	---	--	:23
D-211	12,979	122%	---	--	:31
D-225	7,984	87%	---	--	1:19
S-601	8,074	82%	---	--	1:21
S-621	9,489	105%	---	--	1: 7

OVERALL-KEY-EFFECTIVENESS	RANK
68%	20 OF 24

LAST-WEEK'S-PERFORMANCE (4 DAYS)

OVERALL-KEY-EFFECTIVENESS	RANK
97%	2 OF 28

KEY-MIN	%-OF-MINIMUM	RANK
302	81%	6 OF 28

FORMAT-EFFECTIVENESS	RANK
120%	5 OF 28